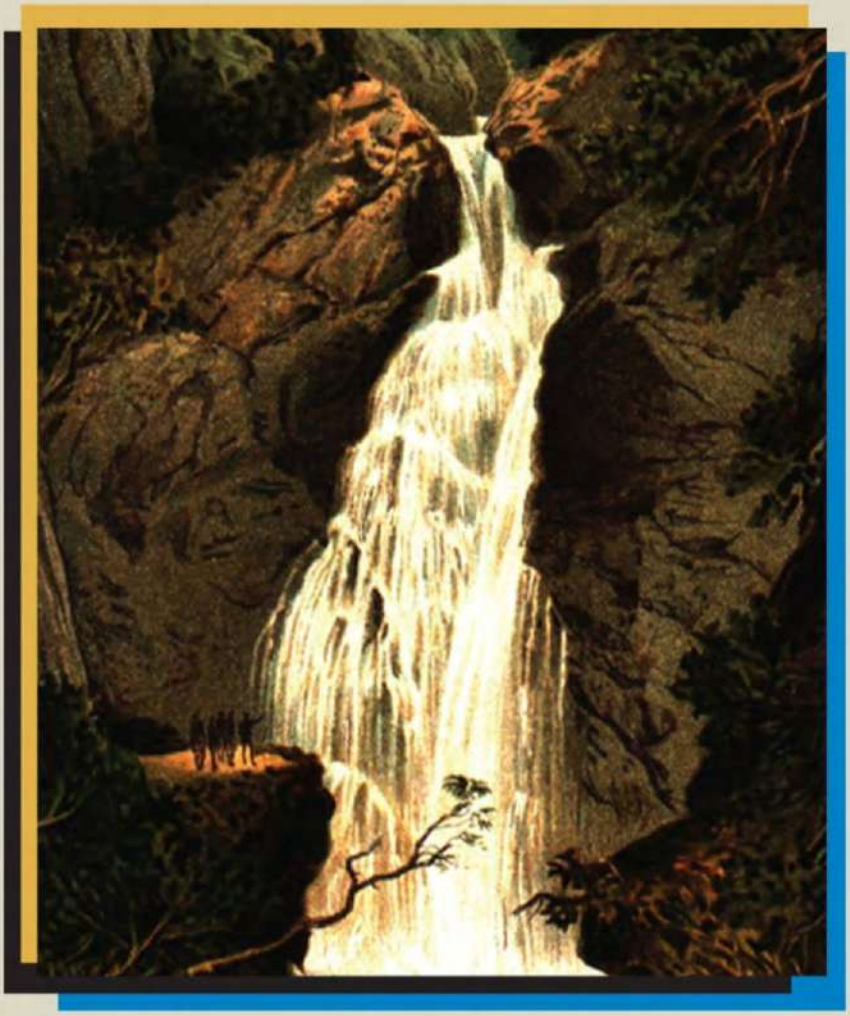
Third Edition

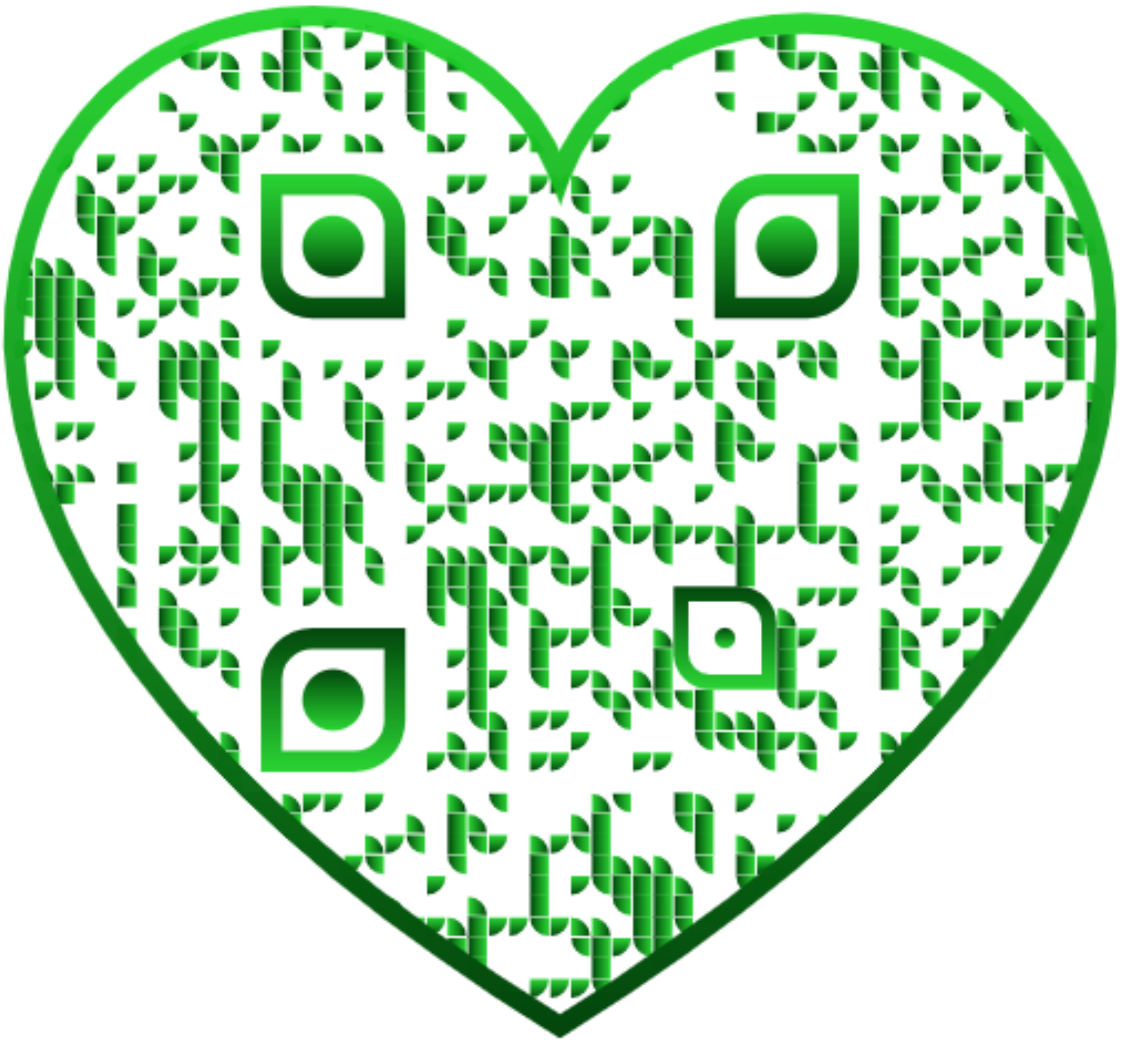
Visual Perception

AN INTRODUCTION



Nicholas J. Wade
and Michael T. Swanston

Psychology Press



Visual Perception

Does the world appear the same to everyone? Does what we know determine what we see? Why do we see the world as we do?

Vision is our most dominant sense. From the light that enters our eyes to the complex cognitive processes that follow, we derive most of our information about what things are, where they are and how they move from our vision.

Visual Perception takes a refreshingly different approach to this enigmatic sense. From the function that vision serves for an active observer to the history of visual perception itself, the third edition has been extensively revised, updated and expanded, while still preserving the essential features of historical context, neurophysiology and independent thought that made the earlier editions so engaging.

Covering the perception of location, motion, object recognition and with up-to-date information on the workings of the visual brain, the third edition looks at how our ideas have been shaped, not just by psychology, but by art, optics, biology and philosophy. The emphasis on understanding vision as a basis for action in the real world has also been expanded to cover seeing representations of all sorts, whether they are pictures or computer-generated displays.

The third edition of *Visual Perception* is a readable, accessible and relevant introduction to the world of perception and will be welcomed by students of visual perception as well as anyone with a general interest in the mysteries and wonder of vision.

Nicholas J. Wade is Emeritus Professor of Psychology at the University of Dundee. He studied at Edinburgh and Monash Universities and at the Max-Planck-Institute for Behavioural Physiology, Seewiesen, Germany. His research is concerned with binocular vision and motion perception, the history of vision research and the interplay between visual science and visual art.

Michael T. Swanston is Emeritus Professor of Psychology at the University of Abertay Dundee. He studied at Cambridge and Dundee Universities and held a Research Fellowship at the University of California Santa Barbara. His research is concerned with spatial representation and motion perception in vision and the perception and measurement of pain.

Visual Perception

An introduction

Third edition

Nicholas J. Wade and
Michael T. Swanston

Psychology Press
Taylor & Francis Group
LONDON AND NEW YORK

Contents

Preface to the first edition ix
Preface to the second edition xi
Preface to the third edition xiii

1 Understanding visual perception **1**
Functions of visual perception 4
Measuring visual perception 10
Visual stimuli for experiments 21
Models of visual perception 23
Reference notes 27

2 Fundamentals of vision **29**
Visual perception and the physical environment 32
Art and representation 43
Development of perception 53
Illusions 60
Variations in vision 70
Reference notes 74

3 Theoretical approaches to vision **78**
Early theories of vision 79
Philosophy 79
Nineteenth-century influences 84
Psychology 87
Twentieth-century developments 89
Twenty-first-century prospects 104
Reference notes 105

4 Optics and the eye **108**
Image formation in the eye 111
Ocular anatomy 114
Limitations of optical performance 117

Measures of optical performance 121
Ocular microanatomy 128
Reference notes 138

5 The visual brain **140**
Visual pathways 141
Visual neurophysiology 145
Neural activity in the visual cortex 154
Midbrain structures associated with vision 159
Neurophysiological interpretations of visual phenomena 161
Visual processing beyond V1 165
Reference notes 168

6 Spatial location **170**
Frames of reference 172
Coordinate systems 178
Visual orientation 181
Visual direction 187
Visual distance 196
Navigation 211
Reference notes 212

7 Spatial motion **214**
Motion phenomena 214
Sources of motion stimulation 228
Frames of reference in motion perception 237
Perceived distance and motion 244
Perceived self motion 246
Reference notes 249

8 Spatial representation **253**
Perceiving objects 253
Perceiving object properties 254
Perceptual constancies 254
Recognising objects 265
Discrimination and generalisation 265
Pictures 267
Reference notes 279

9 Summary and conclusions **282**

References 289
Name index 302
Subject index 307

Preface to the first edition

Vision is our dominant sense. We derive most of our information about the world – about where things are, how they move, and what they are – from the light that enters the eyes and the processing in the brain that follows. These functions are performed by all sighted animals, including ourselves, and yet we still do not understand how. Vision is also the sense about which we know the most, because of the vast amount of empirical research that has been undertaken over the years. This large body of knowledge is celebrated in most of the textbooks that have been written on visual perception; indeed, it can act as a shroud that obscures the purpose of vision from many of those who study it. We feel that textbooks tend to focus too closely on the plethora of phenomena of vision rather than on its function: They frequently reduce vision to a series of headings such as brightness, colour, shape, movement, depth, illusions, and so forth, while remaining blind to the uses to which it is put. We have tried to redress the balance a little in this book. The principal focus is the function that vision serves for an active observer in a three-dimensional environment – we must be able to see where objects are if we are going to behave with respect to them. Thus the perception of location, motion, and object recognition provides the core to the book, and our intention is to make the ideas involved in their study accessible to the reader with no background in psychology. With this in mind we decided deliberately to avoid citing references in the text. This strategy might prove trying for the instructor, but it is hoped that it has the effect of making the book more readable. If it is necessary to qualify every minor point regarding the experimental base of a scientific discipline then it cannot be very securely founded. Another feature we have tried to stress is the historical context in which our present studies are conducted. The history of the study of vision is as long as science itself and many forces have fashioned the conceptual framework in which it operates today – our ideas have been shaped by art, optics, biology, and philosophy as well as by psychology. We need to appreciate these influences if we are to learn from, and avoid repeating, errors from the past.

The other framework that has structured this book is that of three-dimensional space: All our behaviour takes place with respect to it, and since behaviour is guided by perception it is logical that they share the same coordinate systems. In the course of our joint research on space and motion perception we have

developed a scheme for examining the logical stages through which we believe perception passes. This scheme, which involves specifying the frames of reference for extracting visual information, has guided the description of vision presented here. It has involved the introduction of some novel terms in the text, but we believe that this is a small price to pay for maintaining a degree of coherence when dealing with widely disparate topics in vision. It has also resulted in some selectivity of subject matter, but this can be remedied by referring to the general references cited at the end of each chapter.

Science is a social endeavour, and we have benefited from collaboration with many colleagues most particularly with Ross Day, Walter Gogel, Ian Howard, Hiroshi Ono, and Charles de Weert. We also wish to express our thanks to Mark Georgeson who read the manuscript with great care; we have profited from the many insightful comments he made on it. The errors that remain are, of course, our own.

Vision is a subject of enquiry that has fascinated students for centuries and continues to hold our attention. We hope that some of those who read this book will be encouraged to join in the endeavour of broadening our understanding of it.

Nicholas J. Wade and Michael T. Swanston
Dundee

Preface to the second edition

When we began the task of preparing a second edition of this book, we were unsure how much of the original structure and content should be retained. After all, there have been far-reaching developments in visual science during the last decade. Also, we now have experience of students' responses to the book, and the benefit of views from many colleagues. As a result there have been considerable changes, and many sections have been rewritten. However, the essential structure of the book is unchanged. It is still based on a functional approach to vision as a basis for action, and the core chapters are those concerned with perceiving the location, motion and identity of objects in the environment. The introductory chapter has been rewritten and expanded, as has the following chapter on the historical roots of modern concepts about how we see. The chapters on location and motion retain the same emphasis on defining the frames of reference within which object characteristics are represented, but we have sought to make the development of the ideas clearer and more accessible to students. The conceptual model that informed our analysis of space and motion has itself evolved in the last decade, and shadows of these developments are cast in the core of the book and in the Summary and Conclusions. The chapter on recognition has been radically revised, and we have included sections on the major theoretical approaches to this topic. Nevertheless, we have retained the original basis for this account, which is the definition of the representations on which object recognition can be based and how this can be derived from visual and other sensory stimulation. A new chapter on the perception of visual representations such as pictures has been added. This includes a section concerned with the perception of computer-generated images, which is an increasingly pervasive aspect of visual experience.

Much of what was said in the Preface to the first edition is still entirely relevant to the second edition. However, we have decided to introduce a limited number of references in the body of the text, and we hope that this will help students with the critical transition to learning from a wider range of sources. We hope that the book will normally be read as a whole, but it is also possible to use various chapters independently or in groups. Chapter 2, 'The Heritage', can be read on its own, or omitted if historical issues are not of particular concern. Clearly, we would argue that this is not a strategy likely to lead to a balanced perspective on modern achievements, but if time is short this may be a sensible strategy.

Chapters 1, 4, 5 and 6 are the core of the book, and these constitute a minimal selection. Our experience, no doubt shared by others, is that students will probably be drawn to read just those chapters that they have been told to omit, and that this often leads to the most useful subsequent discussion. We have tried throughout the book to reflect the intrinsic fascination of the subject matter, and its relevance for everyday activity and experience. Although there are many illustrations, there are very few of the pedagogical devices of modern textbooks, such as discussion boxes, quizzes and the like. We have eschewed these for the sake of narrative continuity and because we are unsure of their real worth. We hope that the book will demand effort to achieve understanding. This effort will be repaid if it facilitates analysis and discussion in the tutorial setting and a desire to learn more about vision.

We are grateful to a number of people for helpful comments on the first edition, and to very careful and thoughtful readers of a draft of this second edition, including Johan Wagemans, John Harris, Tom Troscianko and Iain Gilchrist. Any errors and infelicities that remain are of course our own responsibility, but they are certainly fewer than would otherwise have been the case. Our families have continued to show the unqualified support and tolerance of our collaboration on this and other endeavours that has made the work possible, and we are glad to have this opportunity to say how grateful we are.

Nicholas J. Wade and Michael T. Swanston
Dundee

Preface to the third edition

We hadn't anticipated writing a third edition of our book and therefore also its preface. However, it has proved to be a rewarding experience and an opportunity to look back at and evaluate the many changes that have taken place in the past two decades. Over this period there have been many remarkable technological developments in visual science – computer-controlled displays as the ubiquitous stimulus sources, enormous advances in neuroimaging, precise measurement of a moving observer's eye movements and a plethora of novel techniques for measuring the manifestations of perception. We have observed their birth and subsequent maturation with fascination. One aspect of these changes has been an even greater reliance on the use of pictorial stimuli to investigate aspects of spatial vision. The number of books dealing with vision has also increased markedly. With so many novelties to nurture, it might seem strange that we have reverted more fully to the approach we adopted in the first edition. We have argued strongly against the equation of pictorial representations of objects with the objects themselves, but this is a feature almost demanded by developments in visual technology. Thus, that which is seen has not been transformed to the same extent as the technology supporting its study. The fundamental problems remain much the same and they remain unresolved. The approach we adopted at first still seems sound to us: perception occurs in a three-dimensional environment and the functions it serves are those of determining where we and objects are, whether one or both have moved, and what the objects are that surround us. That is, the chapters on spatial location, motion and recognition remain at the heart of the book, and the earlier chapters provide the background for appreciating the functional approach adopted. These chapters are linked by a model of spatial vision that is based on frames of reference: the goal of perception is to achieve an internal representation that matches the dimensions of external space. Such a model implicitly involves the behaviours of the perceiver as well as the responses to external stimulation. Indeed, our initial emphasis on perception and action has been reflected in the growing number of studies involving objects in space rather than simulations of them, as in studies of eye movements when navigating through space. Accordingly, the skeleton of the book (chapters 6-8) is unchanged although the flesh upon it has grown in some parts and withered in others.

General issues concerning the ways in which visual perception is investigated remain the substance of chapter 1, whereas the previous chapter 2 is now divided into two: the fundamental and the theoretical aspects of vision. Similarly, the previous chapter 3 (light and the eye) is now separated into considerations of optics and ocular anatomy (chapter 4) and the visual brain (chapter 5). There is no longer a chapter on representation and vision (chapter 7 of the second edition) because much of this material has been incorporated in earlier chapters. Despite our criticisms concerning the use of pictorial stimuli in experiments on vision, pictures can serve an important role in the transmission of ideas about vision. Many of the illustrations in the early chapters (particularly 1-3) are novel and they involve 'perceptual portraits' of those who have added to our understanding of vision. Although they are printed in monochrome, coloured versions can be found at http://neuroportraits.eu/ in many instances. We have made a distinction between those who are portrayed in this idiosyncratic style and those who are shown more conventionally: the latter are for those for whom we have little or no knowledge of their actual appearance, which generally refers to classical students of vision. The inclusion of portraits indicates that we believe the historical dimension of vision is an essential ingredient to appreciating contemporary approaches. In pursuance of this belief, we have tried to integrate historical ideas into the specific chapters, rather than having a chapter devoted to them. The line illustrations in the later chapters are similar to those in the earlier editions.

In the first edition we adopted the strategy of presenting the text without references to specific articles or books; reference notes were appended to each chapter. We were encouraged to include citations in the text for the second edition, and we followed this advice by citing essential studies. We have decided to revert to the original strategy in this edition in the belief that the text will be more readable, and the interested reader can follow up the citations given in the reference notes. There should be a bedrock of evidence about vision that can be summoned without qualification and no introductory reader wishes to be confronted with constant queries about the phenomena and interpretations described. A full list of references is included in the bibliography, and these consist almost entirely of printed works.

We have been assisted by many people in writing this and the earlier editions. We would particularly like to acknowledge Hiro Ono, Charles de Weert, Alan Wilkes, Tom Troscianko and Ben Tatler. The manuscript for the current edition was read by Johan Wagemans, Patrick Green, Frans Verstraten and a fourth, anonymous reviewer. All made valuable suggestions, some of which we have followed. Those errors that remain are, of course, our own.

Throughout all the editions we have received constant support from our families to whom the book is dedicated. Like the book itself, they have grown in number appreciably during the process.

Nicholas J. Wade and Michael T. Swanston
Dundee

1 Understanding visual perception

All living things actively respond to the environment in which they exist; indeed, this may be close to a definition of what it means to be alive. We, descendants of the first land-living vertebrates, have a number of specialised senses that have evolved to sustain survival – to make and maintain adaptive contact with the environment. Receptors for sources of energy in the terrestrial environment that have proved beneficial for survival have evolved and become specialised for the needs of each species. Through the action of their senses, living things seek shelter from the elements and other organisms as well as sustenance and sexual partners in order to survive and reproduce. Thus, environmental features that afford protection can be identified and exploited. Sensitivity to the visual, auditory, olfactory, and tactile characteristics of other living things permits identification of foods as well as potential mates or rivals.

By emphasising the visual dimension of our environmental interactions in this book, we do not wish to suggest that vision operates in isolation from other senses. It is clearly insufficient to study vision alone, as all our senses are intimately integrated, and we have a unified perception of external objects. Perception engages all the senses, but we possess the unique ability to express our perceptual experiences through language, which tends to reflect the operation of particular senses. Thus, it might seem from the way we speak about perception as though the senses work in isolation rather than in concert. The vocabulary devoted to each of the senses is not evenly distributed either. Vision has the lion's share of words as well as the work associated with our perceptual experiences.

Human senses are linked to an intricately organised brain, which has evolved to do more than simply identify and act on the necessities for survival. Humans not only use their senses to survive; they also reflect on their experience through language and thought. Through language, we know that our own and other people's perception of the world is generally stable: objects have positions, shapes, and colours that we seem to be conscious of instantly, and we can reach for them or move to where they are, without any apparent effort. We can recognise small differences between objects and we can categorise or name them appropriately. Some objects, like food, will be particularly significant for our survival, and we learn how to recognise them from all sorts of positions. Perceiving the location of objects and recognising them is achieved regardless of whether we, or the objects

themselves, are stationary or moving. Clearly, there is some process within us that gives rise to our experience of visual perception and the behaviour that it guides. It is not surprising that throughout history students of the senses have found it fascinating to speculate what this might be.

What we call 'perception' is an experience that normally results from stimulation of the senses by the environment. Perceptions can also arise through hallucination – through dreaming, abnormal brain function or the use of drugs. However, our aim is to describe the characteristics of perception in its normal mode of operation, which has evolved to support interaction with the world around us. Perception can be examined in a variety of ways, such as describing our experiences verbally or measuring them by psychophysical experiment; it can also be related to the processes in the nervous system that accompany such experiences. Nerve impulses in the sensory nerves are relayed to the brain and behaviour, including speech, usually results from this sequence of events. Traditionally the only records of action of the senses were provided by behaviour but now a wide range of indirect physiological measures can augment them. In humans the range of behaviours is broad and includes describing the experiences initiated by sensory stimulation and the links it might have with previous stimulations. We refer to these as observations and we associate them with verbal descriptions. Observations provide the bedrock of perception and of other actions of the brain. Historical records of observation precede records of their verbal descriptions; that is, the products of art precede those of writing, by many thousands of years. Relatively little is known about the origins of visual art: examples of marks made on tools and cave walls have been dated to tens of thousands of years ago, but we do not know when such activities began. We assume that the cave-dwellers saw their world in much the same way as we see ours. It is not perception that has changed, but the world in which it takes place.

While we perceive the world around us with alacrity and ease, we have no direct knowledge of how this experience comes about. In fact, it can often be hard to believe that there is any mechanism involved in perception at all; for most people, most of the time, perceptions are simply 'given' as facts about the world that are obviously correct. Perception is not only a basic psychological process, but also a very remarkable one. Its success in providing us with accurate information about the characteristics of the world around us is an index of its power, because there are relatively few situations in which it is seriously in error. A perceptual process that gave rise to subjective experiences grossly different from physical reality would make survival virtually impossible. The function of perception is not to furnish us with subjective impressions of our surroundings and the significant objects in the environment. Rather, it is to provide an effective platform for action. Perceptions and actions take place in a three-dimensional environment and so it is necessary for spatial perception to share the three-dimensional coordinate system in which behaviour occurs.

Visual perception can be studied in many ways. We have chosen a particular approach to understanding our interaction with the world through perception, by emphasising the means by which we achieve correspondence between the

external environment and its internal representation. This approach places special emphasis on the perception of space and spatial relationships (within which we include the perception of motion) as a platform for effective action. This is not to underestimate the significance of the perception of other characteristics of the world such as colour, but it is clear that spatial perception is a consistent requirement for action, whereas colour perception may or may not be. A crucial element of this approach to spatial vision is defining – rather than ignoring or assuming – the frame of reference within which sensory information is represented. When discussing any aspect of space or spatial perception it is necessary to define the frame of reference relative to which measures are made. This is commonplace for physicists, who are accustomed to defining the framework within which any measurement is expressed. Surprisingly, it is less common among perceptual psychologists, even though they are dealing with a system that has detectors in a moving eye, two eyes that can move with respect to one another, and eyes that can be displaced when the head or body moves. Information about where the image of an object falls on the retina could alter because the object moves, because the eyes move in the head, because the head moves with respect to the body or because the body moves in space. Any of these changes, alone or in combination, will result in the displacement of an image on the retina. Generally we perceive the world while there are complex patterns of eye, head and body motion, but our perception of objects is not disturbed. To understand how this might be achieved, we need to understand how frames of reference can be specified, and how they can apply to visual perception. It is this functional path that we will pursue and it is this aspect that we believe is neglected in many books on vision.

There are many excellent texts describing the eye, how it passes signals to the brain and how its many areas refine the physiological processing further. Moreover, these books describe, and often illustrate beautifully, phenomena that can be perceived. These tend to be presented in chapters with titles like brightness, colour, size, shape, depth, motion, and illusions, and they tend to give the impression that these are isolated aspects of visual perception. We have tried to avoid such categories by focusing on the functions of vision in a three-dimensional environment – determining where objects are, whether they are moving and what their identities are. We believe that our emphasis on function rather than phenomenology provides a stable basis for understanding visual perception and for appreciating the importance of historical achievements.

This task is in some ways rendered paradoxical by the fact that the medium we are using to present these ideas lacks some of the properties we wish to accentuate. It is tempting to assume that perception can be understood by studying pictures of solid objects rather than solid objects themselves, but we believe that this is one temptation that should be resolutely eschewed. The pages of a book are flat, as are the pictorial images on them. On the other hand, the book itself is an object in space and we shall refer to it as a means of illustrating issues about frames of reference in spatial perception. In this context, the term ‘book’ could refer to a printed text, an e-reader or any other surface on which text is presented.

A further potential impediment to our purpose is that we are using words to describe the processes of perception. This can encourage the impression that perception itself is like some internal dialogue about the world, and that perception is therefore idiosyncratic. Since people describe the world slightly differently we might think that they perceive it differently, too. We know that this can happen in the case of colour, but less so for the perception of objects. We have tried to discourage such thoughts by employing the assistance of a metaphorical guide dog for a blind person. Clearly such dogs can be trained to sustain the survival of blind people by avoiding obstacles, seeking selected sites, indicating impending danger and even crossing roads. The achievements of such guide dogs are remarkable. They can discriminate between objects in a similar way to us, but they cannot name them. No one would attribute their perceptual abilities to linguistic ability and yet they perform tasks that are equivalent to those of humans. Do we need language to understand human perception?

Another unusual aspect of the book is that there are relatively few references given in the text and, indeed, relatively few experiments that are described in detail. The reason for this is that we have sought to describe aspects of perception on which there is general consensus, like the perceptual constancies, but nonetheless aspects that still pose considerable theoretical challenges. It is in such contexts that we have found the notion of frames of references particularly useful.

This chapter provides an overview of central issues in the study of visual perception, many of which will be discussed in more detail in later chapters. It is important to understand the functions that any visual system must perform if there is to be coordinated, effective action, and the problems of devising explanations for how this comes about. If perception is to be explained, appropriate measurements of its characteristics must be obtained, and related to the information potentially available from the physical environment. Each of these issues contributes to the general framework of ideas that guides the investigation of vision.

Functions of visual perception

We all enjoy contemplating the experiences provided by our senses, and much of our language is associated with describing them. In human cultures considerable effort is devoted to enhancing perceptual experiences by decorating our bodies and our surroundings and by producing artefacts (like pictures) to stimulate the senses and to channel our contemplations. With so much emphasis on extending our perceptual experiences it is tempting to think of their function as enabling us to enjoy and describe them. In evolutionary terms the function of perception is much more mundane – it is to enable us to interact with the objects in the world surrounding us. More specifically, we use our perceptions to guide our behaviour. We use vision to determine the location of objects with respect to us, so that we can approach them, grasp them, cast them aside, or avoid them as appropriate for our survival. Some objects, like food, will be particularly significant for our sustenance, and we learn how to recognise them from all sorts of positions. Perceiving the location of objects and recognising them is achieved when we are

still or moving, or if the objects themselves move. Accordingly, we need to be able to distinguish between static and moving objects whether we ourselves are static or moving.

Action and recognition

As we have emphasised above, vision does not occur in isolation, although we often describe it as though it does. It is integrated with the other senses and also with the movements we make. The actions we perform are based on vision and vision, in turn, is influenced by the actions we perform. This can be illustrated in terms of what you are doing while reading this text, whether from a traditional book, an e-reader tablet or a computer monitor. The surface you are reading from is likely to be in a fixed position and appear stationary. This is the case even though the reading surface and the text on it are moving over your eyes as a consequence of eye movements. We now know quite a lot about the ways the eyes move when reading text and also when viewing scenes: they tend to remain relatively still for a few tenths of a second (called fixation) and then flick rapidly (or saccade) to a new location, with this sequence repeated about three times every second. Despite these jerky movements of the text over your retinas the experience is of a stable surface with stable text. For this to happen there must be an intricate internal integration between the pattern of stimulation on the retinas and the signals for contracting the eye muscles. The situation becomes even more complex when we take gross movements of the body into account; the head can move with respect to the trunk and the whole body can locomote through space.

Thus, at this basic level of perception, action is a fundamental determinant of vision. All observer movements may displace the patterns of light projected on the retinas, irrespective of the location and motion of the objects in the environment that give rise to them. This issue will be discussed in greater detail in later chapters, where we will identify the frames of reference with respect to which object characteristics are successively represented. For example, the retina itself sets the frame of reference for any information deriving from its light detectors and is subject to perturbations from eye and body movements. To be useful for guiding action by a human or a guide dog, this potentially ambiguous information must ultimately be represented in the same frame of reference as that within which action takes place; that is, the three-dimensional environment in which we live.

The complexities of this seemingly straightforward process can be illustrated with an example of a task we perform successfully every day – crossing the road. We will also enlist the cooperation of our metaphorical guide dog in order to contrast its performance and perception with ours. When we wish to cross a road we walk up to the kerbside, look in both directions and cross if it looks safe to do so. But what does this apparently simple sequence entail, and how can it be explained? Crossing the road itself requires coordinated activities in many muscle groups that control balance and locomotion. However, before the walking is initiated a whole series of decisions has to be made on the basis of perceptual information. How far away is the kerb and in what direction is it? How wide

is the road and how much time will be required to cross it? Is there a vehicle approaching? What size is it? How far away is it? How fast is it travelling? How long will it take to reach here? What type of vehicle is it? Each of these questions is directed to some aspect of the three-dimensional environment and our relation to it. The consequences of making wrong decisions for any one of these questions can be life-threatening.

How far away is the kerb and in what direction is it? To answer these questions it is necessary to determine where the pavement stops and where the road starts. This could be done on the basis of some markings (e.g., white lines) which we have learned to associate with roads, or on a difference in the perceived level of the two surfaces. Some theorists consider that we can derive information regarding the orientations of surfaces in the environment without recourse to knowledge about the nature and purpose of objects. That is, the information available in the pattern of stimulation is sufficient to specify the layout of space with respect to the viewpoint of the observer, and assumptions about the nature of the world are not necessary. On the other hand, perception can be likened to thinking and other high level cognitive processes like reasoning and problem-solving. Such cognitive theorists argue that we use our knowledge of roads and their features to recognise the kerbside, and that we make use of a 'knowledge-base' to find and identify objects in the visual scene. Thus what we see is a logical, rational interpretation of sensory information on the basis of stored knowledge, in terms of possible real events. It is certainly not necessary to have verbal labels like 'pavement' and 'road' to perceive the positions and sizes of the actual objects, as guide dogs for blind people carry out such discriminations with remarkable skill. In the discussion that follows, we will contrast the capabilities of a human and a guide dog in crossing a road, as this helps to indicate how far the perceptual processes involved depend on verbal knowledge.

How wide is the road and how much time will be required to cross it? When the edges of the road have been isolated, by whatever means, some estimate of the distances of both from the observer is required. Distance information can be with respect to the observer or between objects in the visual scene. For example, the near kerb could be perceived as two metres away and the far kerb twenty metres distant, then the width of the road could be determined from the difference between the two estimates. Alternatively, the distance between the two kerbs could be perceived in some relative way, such as 'three cars wide'. The time required to cross the road could then be computed as long as there is some value of walking speed available.

Is there a vehicle approaching? Before this question can be addressed it is necessary to define what a vehicle is. An abstract definition in terms of, say, a metal body with windows supported by four wheels, would not suffice visually, as we rarely see all the wheels on a vehicle at the same time; often we do not see any wheels at all. Moreover, some vehicles have two wheels, others ten, and so on, making any definition either very abstract or very long-winded. It is more fruitful to pursue the aspect of motion, as the vehicles of importance will be moving rather than stationary. Is some moving object on the road? The perceptual

problems involved here are those associated with the principles of grouping, for example the perceptual separation of an object from its background. Such perceptual segregation can be difficult for line drawings, and so it may be all the more difficult for a real scene in which the light intensity differences between an object and its background (which may be other objects) can be minimal or non-existent. Assuming an object is segregated from its background, how is its movement perceived? This might be particularly difficult if the background (other objects) also moves.

How fast is the vehicle travelling? A decision as to whether to cross the road is critically dependent on accurate judgement of the velocity of an approaching vehicle. There are differences of interpretation by theorists depending upon the extent to which perception is thought of as 'data driven' or 'hypothesis driven'. In the former case other perceptual information about the distance of landmarks and size of the vehicle might be used to obtain an estimate of velocity. In the latter, knowledge about the likely velocity and acceleration of familiar vehicles could be accessed. One approach to 'data driven' analyses replaced the concept of a static projection to the eye in favour of transformations of the pattern of stimulation over time. Thus the total pattern of light entering the eye could be considered as an optical array, rather like a camera projection onto a screen. However, the transformations of the optical array are considered more informative. These optical flows often occur as a consequence of the observer's own movements in addition to those due to object motions. This approach incorporates the time dimension in perception, so that all perception becomes motion perception. Optical flows can be produced by motions of objects with respect to the background or by motions of the perceiver, the projective consequences of which will be different. If a vehicle is moving then there will be specific transformations in the optic array: it will occlude parts of the road in the direction it is travelling and disclose other parts behind it. This situation only occurs with object motion. Observer motion would result in simultaneous displacements of all parts of the optic array. An alternative approach would determine that the vehicle was moving by comparing the information available in successive time frames. Speed would then be estimated by computing distance divided by time, in the manner of a physicist. The problem with this is that the moving object needs to be identified in successive instants, before its movement can be computed. This requirement for matching different sources of information about the same object is known as a correspondence problem, and it is a feature of many theoretical accounts that rely on analysis of a series of static images. The guide dog can respond appropriately to vehicles travelling at different speeds, so its perceptual system, like ours, can either solve the correspondence problem for movement, or never has to face it.

How long will the vehicle take to reach here? This question is based on the assumption that we have determined how long we require to walk across the road. The answer is dependent upon the information gathered from the earlier questions: if we have determined the location of the vehicle, as well as its direction and velocity, we can estimate the time available for crossing the road. Here the guide dog is superior to humans, for it could cross the road far more quickly that the

person it is guiding. However, it has been trained to estimate the parameters that apply to humans rather than to dogs. Moreover, guide dogs are more likely to use pedestrian crossings than are normally sighted humans! If the approaching vehicle is very close to the observer then the time before contact can be determined from the optic flow pattern. This optic flow changes in a characteristic way that provides information about the time before a collision takes place.

What type of vehicle is it? This is a question of object recognition. A vehicle can be specified as having a certain size, shape and motion at a given distance, but it might not be identified as belonging to a particular category. The categories can be general like 'motor car' or specific like 'red 2012 Ford Focus'. Object recognition involves two aspects, discrimination and naming. The first is essential whereas the second is not: the guide dog will perform the first but not the second. We perform many discriminations for which we have no verbal categories. For example, many people can differentiate between species of fish without being able to name them, although some are able to do both. It is the naming aspect that is characteristically human, and so tends to be investigated at the expense of discrimination. It also creates the impression of some implicit discrimination – some comparison of the presently available instance with a mental model of other instances. Here we have entered the cognitive domain, and we are dealing with processes that need not directly influence behaviour. The guide dog can discriminate between a Ford Focus and a BMW Vision, and its behaviour can be suitably influenced, for instance if the Focus belonged to its former trainer. Discriminations can be made on the basis of the perceived surface characteristics (size, orientation, and texture) of objects, and that can be sufficient to guide our actions.

The approaching vehicle could be a toy bus nearby or a normal bus at a distance. According to their sizes and distances from the observer it would be possible for them both to fill equivalent areas of the visual field, that is, to subtend the same angle at the eye. This comparison seems ludicrous in the context of buses on roads, and so it is, because of the vast amount of other information that would distinguish between the two in the real world. Normally we can see other objects too, and most particularly we can usually see the surfaces supporting the objects. Along with other sources of information, called cues to distance, this would allow the ambiguity to be resolved. Alternatively, knowledge about roads, and the likely objects to be seen on them, could bias the observer towards seeing full-sized vehicles; this would constitute a high-level cognitive explanation. James J. Gibson (Figure 1.1) argued that there is no problem of ambiguity in the real world; rather the pattern of stimulation specifies the sizes and distances of objects. For example the amount of surface texture obscured by an object in contact with the ground remains constant at differing distances, and this invariant feature could determine the perception of constant size despite variations in optical projection. The guide dog can respond appropriately to vehicles of different sizes and at different distances, but this does not rule out any particular explanation. The dog would be capable of using either size and distance information, or optical expansion, and would have extensive experience of vehicles on roads.

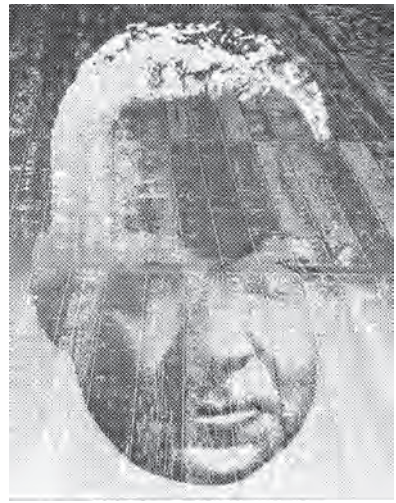


Figure 1.1 **Texture gradients** This portrayal of James Jerome Gibson (1904–1979) is an example of a perceptual portrait, and many more will be presented below. The aim of the portraits is to combine the facial features with a motif that related to the life and work of the person portrayed. In this case, features of Gibson's face and the background are both provided by the same texture gradient of dots but its polarity is reversed vertically: the face is graded from light above to dark below, whereas the background is the reverse. Superimposed on both patterns is a further, ecologically sound, texture gradient – the planks of a pier receding into waters near those lapping Cornell University, where he worked for over thirty years.

Species differences

Guide dogs are selected because of their temperament, their susceptibility to training, and their vision. The fact that they are successful in guiding visually impaired people indicates that they can detect and discriminate aspects of the world in a similar way to sighted humans. However, there are differences in the vision of humans and dogs as well as similarities. The structure of the canine eye is similar to that of humans: it is a single, image-forming system. That is, light reflected from external objects passes through the transparent optical structures of the eye to focus on the light sensitive cells on the rear surface, called the retina. This is only one type of light sensitive organ in the animal kingdom, and there are many other eyes that have evolved to serve the needs of different species. Within vertebrate eyes there are many minor variations in both structure and function. For example, dogs can respond to low levels of light that would be impossible for humans to see. On the other hand, humans can discriminate colour differences that dogs cannot. In large part, this is due to the retinal receptors and their distribution in each species.

The senses which respond to external events play the major role in perception, and species of animals differ greatly in the types of physical energy that they can detect. For example, high-frequency sound waves are employed by bats and

dolphins for recognising objects and for navigation. Certain species of snakes are responsive to the infra-red emissions of warm-blooded prey, and bees to ultra-violet, so that the appearance of flowers to them is very different from that which we experience. In all cases, the aspects of environmental energy that can be selectively detected by animals are those that favour survival. Despite these differences within and between senses the external environmental energy is transformed into a language that can be used by their central nervous systems, whatever their complexity – nerve impulses.

Spatial vision, the ability to discriminate detail in objects and to extract the separations in depth between them, also differs considerably across species. Some of the variation is due to the locations of the eyes and others are a consequence of the particular receptors in the retina. For example, the visual field of dogs is far greater than that of humans: their lateral view is about 240 deg whereas ours is about 200 deg, but we have a larger area that is shared by both eyes. This is because dogs' eyes are not directed straight ahead as is the case for human eyes. Various tests are used to determine the limits of spatial vision, and visual acuity in humans is about four times better than that for dogs. What we see sharply can be a fuzzy blur for dogs. Despite these differences, they can be trained to be vital aids to humans with restricted vision. Certain primates (like macaque monkeys) have very similar spatial and colour vision to humans, and so they are increasingly employed as subjects in perceptual experiments, particularly when aspects of the underlying physiology of vision are investigated.

Measuring visual perception

While each of us experiences perception of the world, obtaining useful measurements of this experience can be very difficult. The attempt to communicate subjective experience to other people has fascinated and frustrated writers, painters and other artists for centuries. It is only through communication that we can convey what we see, hear, taste, smell, and feel. Our language is replete with words relating to the senses, but this abundance can provide problems for understanding perception, and alternative procedures have been developed that rely less on language and more on other responses. The guide dog for a blind person clearly perceives the space around it in much the same way that we do but it does not have recourse to language. We make inferences regarding its range of perception on the basis of experiments that neither use nor require language. We will examine the many methods for measuring perception generally and visual perception in particular in this section.

One of the standard procedures developed for scientific enquiry into complex natural phenomena is to reduce them to simplified situations in which relevant features can be isolated and controlled. In the case of colour vision, we would want to ensure that a stimulus produced light of a known wavelength and intensity. We might wish to control the size of the stimulus, and perhaps its shape. Other sources of light would need to be eliminated, unless the effect of these was to be specifically studied. Such control over stimulation would generally require

laboratory conditions, and usually special apparatus as well. Experiments carried out in this way can provide unambiguous and detailed measurements of visual performance. This approach has been very influential in studies of perception, which often involve visual environments so restricted as to be far removed from natural visual experience. Measurements of perception are obviously easier to obtain when the perceptual experience is itself very simple; for example, sitting in a dark room and pressing a button if and when a single faint light source is seen. The measurements obtained in an experiment are used to infer the nature of perceptual processes, since we can never measure perception directly. It might be thought that this makes the study of perception subjective and indirect, but the situation is not fundamentally different from, for example, an investigation of biochemical processes. Here, too, measurements are used to infer the underlying chemical reactions, which cannot be directly observed. Probably the opportunity for error due to uncontrolled factors is greater with perception, since the system being studied is a good deal more complex, and responsive to a wider range of influences.

Verbal description

The simplest approach is to ask someone to describe their experience, and to draw conclusions from their reports. Descriptions of the same scene or event can be compared across observers, and it may be possible to classify the verbal reports to give some degree of quantification. In principle, free description of experience offers potentially the richest source of information, since language is the most flexible means of communication we have. For many centuries, philosophers and others interested in perception, relied upon verbal description as the only means of obtaining data for analysis. Although perceptual experience is subjective, we are able to communicate quite effectively with other people regarding the nature of the world around us; disagreements about experience are much less likely than agreements. While language is a powerful means of communication, it is nevertheless restrictive; in the limit, only those experiences for which we have words can be described.

Reliance on verbal descriptions has not always clarified our understanding of perception, and this can be illustrated in the context of colour vision. Suppose two people look at a piece of coloured paper: one says that it is red and the other that it is black. Can the same object have two different colours simultaneously? It does not seem likely for a piece of paper, so why are different colours reported? One possibility is that the two people are having the same visual experience, but they are describing it differently – the verbal labels attached to the same experience differ. This could be tested by asking them to describe two other coloured papers; suppose they agree in calling them green and blue, respectively. The problem becomes complicated by this because some colour names correspond and others do not. It would be odd to say that they were each having the same experience for all three colours but only one was described differently. Another possibility is that one of the two individuals is colour defective, that is, one does not have

colour experiences like the majority of the population. If so, which one is colour defective? A straightforward way of determining this would be to ask a number of other people – a sample of the population – to describe the same coloured papers, and to note their descriptions. If most of the sample say that the initial paper is red, then it would indicate which of the two people is colour defective. It seems more reasonable to account for individual differences in perception in terms of variations in the mechanisms of perception rather than changes in the world.

Some perceptual experiences vary in intensity rather like the variations in stimulus intensity, so that the question "How much?" can be applied to them. For example, length is such a dimension: variations in length can be measured physically (say in cm) and its perceived intensity can be given a value – we can apply a number (often in similar units) to the length of an object. There are many such dimensions of sensory intensity, like loudness, brightness, duration, and temperature. Such experiences are distinguished from others for which the dimension of intensity is inappropriate, even though there is a systematic change in the underlying stimulus dimension. For example, variations in the wavelength of light do not result in variations in intensity, but in colour. Changing the wavelength of light from, say, 550 to 650 nm produces a change in experience from green to red. The appropriate question here is "What kind?" rather than "How much?"

Psychophysical methods

Verbal descriptions generally provide qualitative indices of what we perceive. In order to study perception in more detail, quantitative measures are required; that is, measures to which numbers can be assigned. With quantitative measures experimental manipulations become possible and hypotheses about the nature of phenomena can be tested. For example, the phenomenology of colour has been very successful in highlighting the aspects of colour vision that require explanation, like colour naming, primary colours, and contrasts between colours. However, in order to study any of these phenomena in more detail it is necessary to relate aspects of the stimulus (like its wavelength) to features of the response. Stimulus definition is in the domain of physics and the response to stimulation is the province of psychology – so relating one to the other was called psychophysics by Gustav Theodor Fechner (Figure 1.2) in his book with this title, which was published in 1860. For colour many possible response measures could be taken: detecting the presence of one colour amongst other similar ones; matching the colour of one stimulus to that of another; judging whether pairs of coloured stimuli are the same or different. The physical aspect of psychophysics is the measurement and control of the stimulus, and the psychological feature is the measurement and control of the response. It is generally the case that very simple responses are required, like Yes/No, rather than detailed descriptions of experience.

Psychophysical procedures have been employed to determine the limits or thresholds of perception. Detection thresholds are concerned with detecting the presence of a stimulus of low intensity – e.g., is a light on or off? Difference

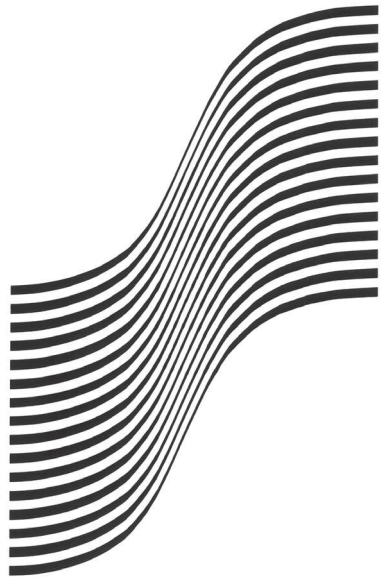


Figure 1.2 **Psychometric function** Gustav Theodor Fechner (1801–1887) is shown in a series of curves which follow the psychometric function. Fechner's portrait is to be found in the area of uncertainty. In this and in some other perceptual portraits when the face is not immediately seen, viewing the picture from afar will assist in seeing the otherwise hidden portrait.

thresholds are of much greater use: they concern discrimination (the detection of a difference) between two stimuli – e.g., is one light brighter or dimmer than another? Typically, a standard stimulus is presented against which a comparison is judged. The observer indicates whether the comparison is greater or less than the standard on some dimension (e.g. brightness). When they are of similar intensity the judgements can be very difficult, and the observer can make different decisions when the same stimuli are presented on separate trials. Because there is some uncertainty in the observer's decision, and therefore variability in the response, it is essential to measure the same conditions many times. Accordingly, the threshold is a statistical concept; it does not signify an abrupt change between not detecting then detecting a difference between two stimuli. Rather it is an arbitrarily defined point in the gradual transition between these states.

There are various psychophysical methods that can be applied to measure perception. Most are based on the methods developed in the nineteenth century by Ernst Heinrich Weber (Figure 1.3) and Fechner in Leipzig. The classical methods are now called the method of limits, the method of constant stimulus presentation, and the method of adjustment. The method of limits involves presenting stimuli in an ascending or descending sequence; the sequence stops when the subject's response changes – e.g. from Yes to No, or vice versa. The stimuli change in intensity by fixed steps and they start from a value that is either easily detectable or

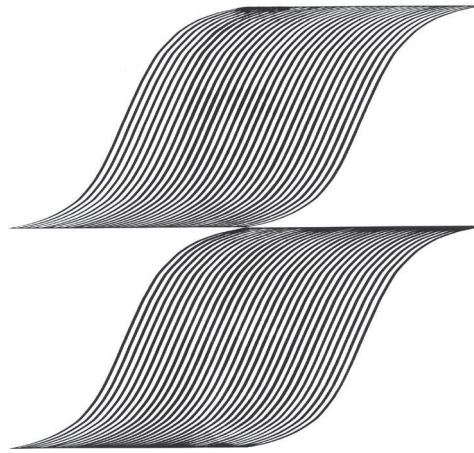


Figure 1.3 **Weber fractions** The portrait of Ernst Heinrich Weber (1795–1878) is carried in a motif representing a sequence of psychometric functions – the curves relating the detection of stimulus differences to stimulus intensity. He is portrayed twice close to the upper and lower thresholds for discrimination and it is from such curves that the Weber fraction can be determined.

impossible to detect. Another classical psychophysical method, which has many similarities, is the method of constant stimulus presentation. In both methods stimuli are presented discretely; that is, the stimulus is presented at a single value and a judgement of it is required. Typically a forced-choice judgement is required – either Yes or No. In the method of limits the next stimulus value depends on the last response, whereas in the method of constant stimulus presentation the sequence of stimuli is random. The range of stimuli to be examined is determined in advance, and each stimulus value is presented many times, so that a probability of detection or discrimination can be calculated. In order to prevent the observer from responding Yes to every trial, catch trials (in which no stimulus is presented) are also given. If observers respond Yes to these then they are probably guessing rather than detecting the stimulus. The method of constant stimulus presentation provides data that can be represented graphically; the graphs can be either in terms of the obtained values or of some smoothed function derived from statistical curve-fitting procedures. The curve-fitting procedure produces an S-shaped relationship which is called the psychometric function (see Figure 1.4). For detection, where observers indicate only Yes or No, the 50 per cent point on the psychometric function defines the threshold. For discrimination, where observers state whether the comparison stimulus is greater than or less than the standard stimulus, the 50 per cent point defines the point of subjective equality (PSE); for some tasks (like visual illusions) this can differ systematically from the point of objective equality (POE). The stimulus range between the 25 per cent and 75 per cent points on the function is referred to as the interval of uncertainty (IU), and the difference threshold is IU/2; the difference threshold is also called the just noticeable difference (jnd).

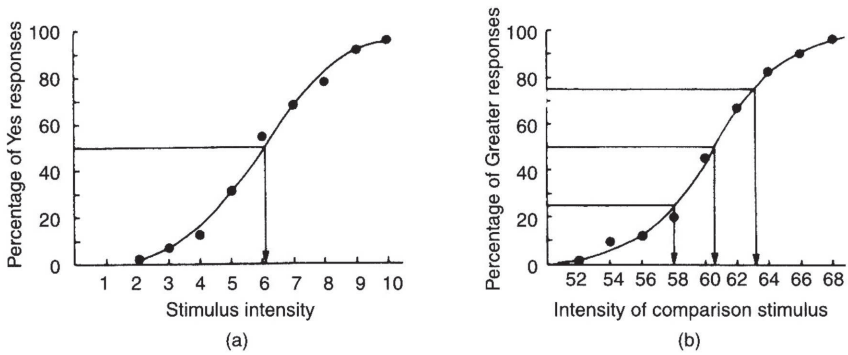


Figure 1.4 Schematic representation of psychometric functions (a) for detection and (b) for discrimination thresholds.

Both the method of constant stimulus presentation and the method of limits have certain disadvantages. The former is very time consuming, and involves presenting many stimuli that are not difficult to judge; it also requires prior knowledge of the likely threshold so that a suitable range of values can be predetermined. The latter can take many presentations to reach the area of uncertainty (where judgements are difficult to make); there are errors of anticipation and errors of habituation; only one point from a sequence (where there is a change in response) is used in computing the threshold or PSE. The method of adjustment generally provides a more efficient means of measuring thresholds. It consists of the observer adjusting the stimulus to find the threshold. Any single adjustment will have an error of measurement associated with it and this can be reduced by taking the average of many judgements rather than one.

Most information can be derived from judgements that are difficult, i.e., those in the area of uncertainty. Modern modifications of the classical methods have tried to find the area of uncertainty with greater efficiency. Staircase methods (sometimes referred to as adaptive testing) have been particularly successful in this regard. They are similar to the method of limits in so far as the value of the stimulus on trial n depends on the response on trial $n - 1$. However, rather than stopping the sequence when a reversal is made, the sequence continues – concentrating the presentations in the region of greatest difficulty. A single staircase is rarely used because an observer could soon learn to alternate responses. Rather, interleaved staircases (two or more) are presented, preferably with random selection between them in order to avoid any anticipation by the subject. There will be some stopping rule to determine the threshold; this is some number of response reversals (like 6 or 8), and the threshold is taken as the mean of these values. An even more efficient version of a staircase procedure is called Parameter Estimation by Sequential Testing (PEST). Not only is the method interactive (adaptive), but the step-size itself varies according the observer's response: each time a response changes the step-size is halved. This

means that large step-sizes can be given initially so that the area of uncertainty is rapidly located. The reversals thereafter become increasingly small, and the stopping rule is defined by reaching a certain (small) step-size. Again, randomly interleaved sequences can be presented so that the subject is not able to predict the stimulus value on a given trial.

The original concept of threshold, in the nineteenth century, was an abrupt change from non-detection to detection or vice versa. As we have seen, it is now considered as a statistical concept, but is the 50 per cent point on the psychometric function a constant value? In the first half of the twentieth century, laboratory experiments to determine absolute thresholds were found not to correspond to the performance of operators in practical situations where the threshold measured depended on the importance of detecting a stimulus. That is, the threshold measured depended upon the motivational state of the observer. Accordingly, a greater appreciation of the decision processes involved in psychophysical judgements led to a novel approach to psychophysics which was called Signal Detection Theory. It involved procedures that could be employed to distinguish between the detectability of a stimulus and the criteria that affect the observer's responses. Signal Detection Theory can be applied when there are considered to be two states of the world for which there are two possible responses. In the context of detection thresholds the two states of the world are referred to as noise (N), which is variable in intensity and always present, and signal + noise (SN); the responses are Yes (there is a signal) and No (there is not). SN and N are presented many times with the stimulus intensity (signal) set at one level. Factors influencing the observer's decision are the signal probability (the proportion of trials on which a signal is present) and the payoff (the rewards and/or punishments associated with a given outcome). Both influence the threshold measured in the classical way, but signal detection theory provides a way of determining detectability independently of the observer's criterion.

Psychophysical scaling

The classical psychophysical methods were described by Fechner not only to provide systematic techniques for measuring thresholds but also to establish a quantitative index of sensory magnitude. That is, he wanted to use the results to scale the intensity of sensations with the precision that scientists apply to scaling physical dimensions, like light intensity, weight or time. In other words, he sought to devise units in which sensory intensity could be measured. The unit he selected was the difference threshold or jnd. Therefore any sensation could, in principle, be measured as so many jnds. But what was the starting point? Here he chose the absolute (or what we would now call the detection) threshold as the zero point on the scale. Weber had earlier found that the value of the jnd increases with the intensity of the stimulus with which it is being compared. This is now called Weber's law and it can be described very simply: $dI/I = k$, where dI is the jnd, I is the stimulus intensity against which a variable is compared, and k is a constant called the Weber fraction.

We can use as an example a task that was studied intensively by Weber and Fechner – comparing lifted weights. Let us assume that the standard weight is 100 gm ($I = 100$) and a range of lighter and heavier ones are compared with it, and that the jnd is 2 gm ($dI = 2$). This is equivalent to saying: 98 gm is judged as lighter than 100 gm 50 per cent of the time and 102 gm is judged as heavier 50 per cent of the time. Then $dI/I = 2/100 = 0.02$. That is, the Weber fraction (k) is 1/50 (0.02). Now suppose that the standard weight is 1000 gm; what weight will be discriminated as different from this? If we know the Weber fraction and the standard we can calculate the jnd: $dI/1000 = 0.02$, therefore $dI = 20$. This is equivalent to saying: 980 gm is judged as lighter than 1000 gm on 50 per cent of the trials and 1020 gm is judged as heavier 50 per cent of the time. The value of the Weber fraction is about 0.02 for heaviness, but it will vary for different types of judgement. For example, the Weber fraction for brightness is about 0.08, for loudness about 0.05, and for line length about 0.03.

Fechner also made a critical assumption – that equal differences in jnd give rise to equal sensation differences. Applying these assumptions to any stimulus continuum results in a curve that rises rapidly and then levels off (called negatively accelerating), it can be represented as a straight line if the magnitude of sensation (S) is plotted against the logarithm of stimulus intensity (I). More generally, the relationship can be expressed as: $S \propto \log I$, which we now call Fechner's law. In words, the magnitude of sensation is proportional to the logarithm of the stimulus intensity.

How can this relationship be tested empirically? It would be very tedious to measure an ascending range of jnds, and Fechner did not believe that sensation could be scaled directly. That is, he did not consider that observers could report the magnitude of sensation. Instead, he used an indirect technique called category scaling. In brief, this involved presenting observers with a wide range of stimuli, and asking them to order some of them into, say, seven categories, so that the differences between categories were subjectively equal. Results using this method generally supported Fechner's law. However, with more research, doubts were cast both on the method of category scaling and on the assumptions Fechner made. Firstly, the validity of Weber's law has been questioned by results from experiments on signal detection. Detection and difference thresholds can be modified by the motivation of observers and by the likelihood of stimuli occurring. Secondly, all jnds do not appear subjectively equal: jnds at the extreme ends of a stimulus dimension do not seem the same as those in the middle region.

Stanley Smith Stevens (Figure 1.5) spent many years developing direct methods of sensory scaling. The most basic of these is called magnitude estimation. As its name suggests, it involves an observer assigning a magnitude – usually a number – to each stimulus as it is presented. Different individuals will give widely different numbers for any given stimulus, but a given individual is generally consistent in the relative values given. There is a simple way of obtaining greater comparability between the magnitude estimates of different observers, and that is by assigning a number to a standard stimulus. This is called a modulus, and it is presented before each judgement. For example, in scaling line length the modulus could be

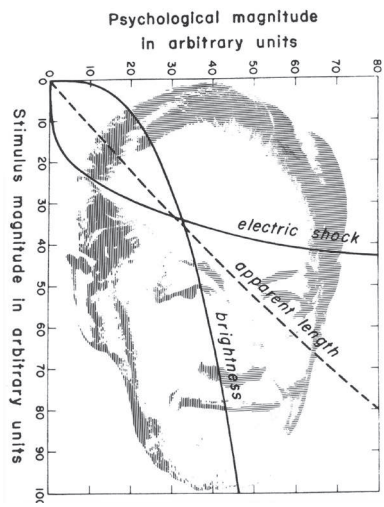


Figure 1.5 ***Power functions*** Stanley Smith Stevens (1906–1973) is shown within the dimensions beloved of psychophysicists – psychological and physical magnitudes – with the former measured by magnitude estimation. The three functions are plotted on linear-linear axes and only one of them (brightness) would conform to Fechner's logarithmic law. Plotting the curves on log-log axes yields straight lines, the gradients of which are the exponents in Stevens' power law. The brightness curve separates vertical and horizontal lines in the portrait of Stevens, one set being slightly brighter than the other. The lines are also of different apparent (and physical) lengths.

a line of a particular length assigned the number 100; other lines could then be presented so that the observer estimates their numerical magnitude in comparison to the modulus.

From magnitude estimation experiments using many other stimulus dimensions (like loudness, brightness, heaviness, duration etc.) Stevens obtained curves that could not be described by Fechner's logarithmic relationship. However, he did find an alternative equation that described all the data from magnitude estimates, and other direct scaling methods. This was of the form $S \propto I^n$ (which is now called Stevens' power law). That is, the magnitude of sensation (S) is proportional to the stimulus intensity (I) raised to the power or exponent n. Stevens' law follows from making a slightly different assumption to that of Fechner, namely, equal ratios of jnd give rise to equal sensation ratios. The value of n varies according to the stimulus dimension examined. For loudness $n = 0.6$, for brightness $n = 0.5$, for heaviness $n = 1.45$ and for duration $n = 1.1$. In the case of apparent length the exponent is 1.0 (i.e., $n = 1.0$) and so a graph of length estimate against the line length should be a straight line with a slope of 1.0. For other dimensions, like brightness, the exponent can be determined by plotting the data on log-log coordinates; the slope of the straight line is the value of the power.

Stevens' law is generally considered to be preferable to Fechner's law because it is supported by data from methods that are less prone to bias than indirect

category scaling. Many different methods of direct sensory scaling (like cross-modality matches and ratio methods) yield similar exponents. There is also growing evidence that the neural responses underlying stimulation might follow a power rather than a logarithmic function.

Physiological indices

Sometimes there may be no overt behavioural indication of perception, and more indirect physiological measures may need to be used. For example, a person may be present at a concert, but give no indication from their behaviour as to whether they are engrossed with the music or bored to the point of sleep. For vision, we usually show by our eye movements that we are observing an event, but this is not invariably the case. However, it may be possible to measure a physiological response, such as a change in brain activity or in the electrochemical state of the retina, using non-invasive techniques. The earliest such measures were the electrical changes recorded from the surface of the scalp, reflecting the activity of many brain cells – the electroencephalogram (EEG). With improvements in amplifying the signals it became possible to record electrical activity associated with single stimuli (if they are presented many times); they are known as evoked potentials or event-related potentials (ERPs). For example, the event might be a checkerboard pattern that is changing phase (black squares replacing white and vice versa) at a given rate (temporal frequency), and the small potentials evoked by each change could be measured. Because the ERPs are so small (often only several millionths of a volt) the events need to be repeated many times and the potentials averaged by computer, so that the background neural noise can be reduced. Studies of ERPs are of most interest when they can be correlated with verbal or behavioural indices of perception, but they may be of value even when the latter are not available. Similarly, in experimental animals measures of the activity of single cells may be obtained, at various points in the visual system; much of our knowledge about visual processing has derived from such studies. Electrical changes in the brain are measured by ERPs whereas the minute magnetic events associated with neural activity can be detected by magnetoencephalography (MEG). The magnetic signals are so small that they can best be detected (using an array of magnetometer coils) from activity close to the surface of the brain.

More indirectly, physiological measures may be taken of functions which are partly determined by perception, such as heart rate or blood pressure. While the occurrence of a physiological response to a stimulus indicates that a corresponding perception may also take place, it is not a guarantee of this. On the other hand, we may have to be careful not to infer the existence of a perception without careful examination of the evidence. A well-known example is that of the ‘red rag to a bull’; bulls in fact cannot discriminate red rags from green or blue rags of the same brightness, and the effect, if any, is probably due to movement.

In recent years novel methods of neural imaging have been developed, like positron emission tomography (PET scans) and magnetic resonance imaging

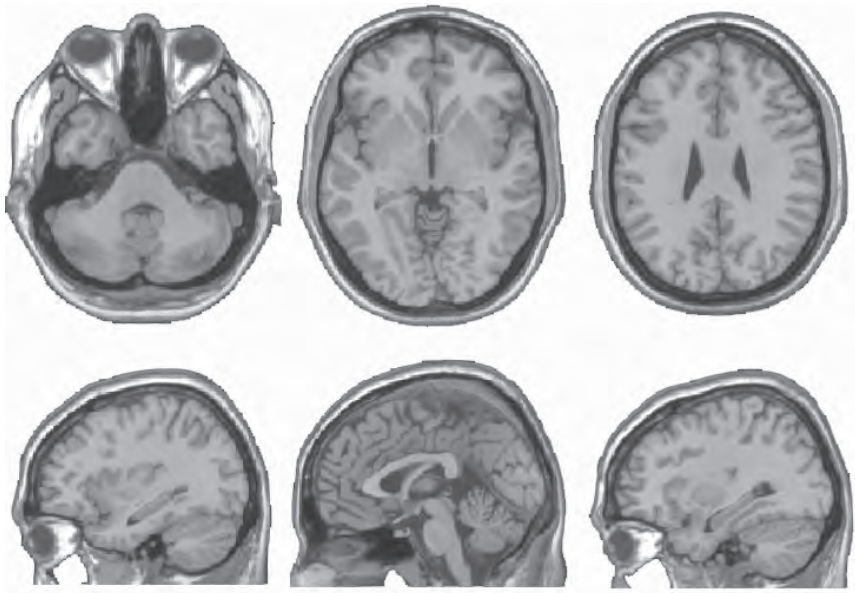


Figure 1.6 MRI brain scans. The sequence of three in the upper half show horizontal sections through the eyes and ascending upwards from that level. Those shown in the lower half are vertical sections through one eye, the median plane and then the other eye.

(MRI). A series of sections derived from MRI scans is shown in Figure 1.6. Both PET scans and MRI rely on computerised tomography which converts small signals from a range of positions into a model of the brain; the images can subsequently be sliced and rotated. They measure activity in the brain and they have proved helpful clinically because the location of lesions or tumours in the brain can be made with greater accuracy. They can also be employed to correlate activity in a variety of brain sites with perception and cognition. When some radioactive substance has been administered and absorbed into the blood the brain areas which are subsequently most active have the highest radioactivity, and this can be detected by PET scans. Thus, the sites of heightened neural activity associated with a particular perceptual task can be determined following suitable computer image manipulation; these areas receive a greater blood flow and therefore provide a stronger signal for the detector. The computer manipulated images can be coloured to signify the regions in which activity has been strongest. Because short-lived, radioactively labelled substances are needed for PET scans there are limits to the measures that can be taken from one person. These constraints do not apply to MRI measures; the observer is placed in a strong magnetic field which aligns atomic particles in the brain cells. Bombarding them with radio waves results in them producing signals, which differ according to tissue type, and the signals can be detected. Functional MRI (fMRI) measures are much more useful for perceptual research, as they are concerned with differences in MRI measured in control and experimental conditions. Cells that are active under the

experimental conditions utilise more oxygen and can be detected and imaged. The spatial resolution of fMRI is superior to that of PET scans. Both these procedures correlate changes in oxygen or nutrient use with brain activity whereas MEG measures magnetic consequences of neural activity itself. The temporal resolution of MEG is very high (around 1 msec) and the spatial resolution is best for neural activity close to the surface of the cortex, and so it is well suited to sensory studies.

The techniques mentioned so far record the activity in the brain associated with some pattern of stimulation. An alternative strategy is to disrupt neural activity in some way. This is achieved with transcranial magnetic stimulation (TMS). A magnetic coil is positioned over a particular area of a subject's head (usually defined by prior MRI measures) and a current is briefly passed through the coil. The magnetic field so produced induces an electrical current in a specific part of the subject's brain. The timing of such TMS is very precise and so it can be applied at known intervals after some visual stimulation has taken place. It is as if the technique produces virtual patients because the disruption is temporary.

All the techniques mentioned have virtues and drawbacks, typically depending on their ability to resolve neural activity in space and time. MEG and ERPs measure events in time very well, but it is difficult to determine precisely where in the brain the signals originated. PET scans and fMRI have the opposite problem; their spatial resolution is good (they can localise areas of activity within the brain) but their temporal resolution is poor (they require long periods, seconds or minutes, to obtain their measures). TMS has good temporal resolution for its application but it can be difficult to determine the spatial localisation; brain currents induced in one area might extend beyond the desired brain location.

Visual stimuli for experiments

The traditional approach to studying perception has been to define the qualities and quantities of physical energy that can be perceived, as for example in the measurement of the least intensity of light that can be reliably detected, or the difference in wavelength between patches of light needed to make them appear discriminably different in colour. These types of measurement define the limits of perception, so they can be used to find out if any perception is possible. They are therefore useful for identifying perceptual defects, as for instance when an individual's discrimination of certain wavelengths is systematically worse than other people's. Ultimately, the purpose of perception is to enable humans and other animals to guide their behaviour in a way that is appropriate to the real environment, whose most pertinent features are the location and nature of objects in three-dimensional space.

Most perceptual research is now conducted with computer-controlled displays, because these provide the most extensive, accurate and precise control of both the stimuli that can be presented and the responses that can be measured. Increasingly powerful computers and graphics systems have replaced many of the classical nineteenth-century instruments, like the stereoscope, stroboscope, and tachistoscope for studying perceptual processes. Devices like these were used for

more than a century to address fundamental questions in perception, and to define the types of answers that could be obtained at the psychological and physiological levels. Present-day research has benefited from the discoveries and theories of the previous era. With the help of computer-controlled displays, stimulus variations that in the past would have been impossible to achieve or required a specialised laboratory, are routine and inexpensive. This can be seen as the culmination of the process by which psychology emerged gradually as an independent discipline because a central branch of it, perception, developed its own methods of both stimulus control and response measurement.

Despite profound technological advances, one characteristic of perceptual research has remained the same. Fundamentally, whether perception is studied with a stereoscope or a computer-controlled display, the intention is to vary a precisely defined stimulus property and to obtain a similarly precise measure of the resulting percept. In this respect, perception is studied by means of the classic scientific method; the properties of a complex event are reduced to a number of simple and measureable variables, all of which are either known or held constant apart from the variable of interest. This approach has yielded a great deal of valuable information. Nonetheless, it inevitably involves testing perception with stimuli that are unlike those encountered in the real environment, for which perceptual systems are attuned through evolution.

Visual stimuli in the environment

Evolution has ensured that perceptual systems are adapted to the needs of organisms, whether for locomotion, foraging for food or finding a mate. In the most general terms, the environment can be thought of as a source of many varieties of potential stimulation; ranging from vibrations in the atmosphere to electromagnetic radiation of various wavelengths and molecules diffused from an evaporating source. Those aspects of the world that we can perceive depend in the first instance on the senses we possess. The world as it appears to us is therefore based on a selective sample of the many forms of energy that are available. One of the developments of modern technology has been the extension of our natural senses to a wider range of environmental events. We can convert x-rays, ultrasound and infra-red to forms that can be directly perceived. Events at a distance far beyond the range of the senses or occurring in a hostile environment can be relayed to an observer. Whatever the sources of information, however, the outcome is perception of a physical environment which has spatial extent and in which objects can be located in space and time, with a particular size and shape. All perceptual systems must provide some such representation, even if it is based on a limited selection of available information, because they would otherwise be of little value for guiding behaviour.

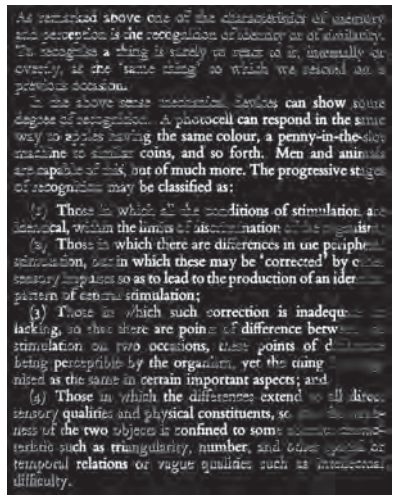
Not surprisingly, it has been argued that we should be cautious about supposing that perception in the real world can be explained by the processes revealed under laboratory conditions. Gibson, in particular, argued that natural perception depends on complex patterns of stimulation, involving active exploration of the

environment by the perceiver. He reacted strongly against the idea, drawn from the physical sciences, that any complex process can be understood by combining a number of simpler component processes. Under restricted laboratory conditions, perceptual processes are not reduced to their elementary components, but rather to an unrepresentative and impoverished form. Thus, Gibson claimed that visual illusions were simply the consequence of looking at scenes with very little information; under natural conditions in the environment we generally do not see illusions because there is plenty of information to tell us about the true sizes, shapes, and colours of objects. Crucially, if observers are prevented from moving then their perceptions will be both unnatural and relatively uninformative. In our view, an adequate account of perception should be able to explain what is seen under any circumstances, whether in a natural environment or in the laboratory.

Models of visual perception

Our understanding of perceptual processes like those outlined in the previous section has very often been shaped by concepts and models drawn from other fields of scientific enquiry and applied as analogies to the working of the brain. Analogies drawn from physics, engineering and computer science (amongst others) have provided ways of describing and interpreting the processes that give rise to perceptual experience. Such analogies may be widely accepted, to the point of being thought self-evident, but their inadequacy has become apparent before long. Clockwork automata of great ingenuity were built in the seventeenth and eighteenth centuries, and these seemed to indicate that living organisms might be thought of in terms of similar types of device, although of much greater complexity. Before the electrochemical nature of nerve impulses was understood, communication in the nervous system was explained by analogy with pneumatic or hydraulic devices, which were in common use. Later, in the nineteenth century, analogies with power and force were widely employed. In the early twentieth century the brain was described in terms of the functions of a telephone exchange, switching messages from one point to another. No doubt all such analogies are useful at the time they are proposed, but it is important to be aware that they are speculations rather than explanations, and that this applies as much to the computer as it does to clockwork. It is simply a measure of our ignorance that we do not know how to characterise the operation of the brain in terms that are independent of analogy with other sorts of mechanism.

At the present time, many of the terms that are used to describe perception, as well as other psychological processes, are drawn from the vocabulary of computer science and electronics. This has been a feature of psychology since the 1940s, and it has resulted in new approaches to long-standing problems. During World War II, devices were created whose operation could be controlled by the consequences of their own actions, the study of which became known as cybernetics, founded by Norbert Wiener (1894–1964). At the same time, the first true general-purpose programmable computers were constructed in the UK and USA. Kenneth Craik (Figure 1.7) and Alan Turing (1912–1954) saw cybernetics



As remarked above one of the characteristics of memory and perception is the recognition of identity or of similarity. To recognise a thing is merely to react to it, mentally or overtly, as the 'same thing' to which we reacted on a previous occasion.

... the above sense-mechanical devices can show some degree of recognition. A photocell can respond in the same way to apples having the same colour, a penny-in-the-slot machine to similar coins, and so forth. Men and animals are capable of this, but of much more. The progressive stages of recognition may be classified as:

(1) Those in which all the conditions of stimulation are identical, within the limits of discrimination of the organism;

(2) Those in which there are differences in the peripheral stimulation, but in which these may be 'corrected' by central sensory processes so as to lead to the production of an identical pattern of central stimulation;

(3) Those in which such correction is inadequate or lacking, so that there are points of difference between the stimulation on two occasions, these points of difference being perceptible by the organism, yet the thing is recognised as the same in certain important aspects; and

(4) Those in which the differences extend to all direct sensory qualities and physical constituents, so that the sameness of the two objects is confined to some abstract characteristic such as triangularity, number, and other spatial or temporal relations or vague qualities such as intensities, difficulty.

Figure 1.7 **Recognition of identity** Kenneth John William Craik (1914–1945) can be recognised in text describing the problems faced by natural and artificial pattern recognisers: "The main parts of such a system are the sensory device, the law relating its controlling effect on the counteracting cause which it produces, and the nature of that cause ... The essential feature of the sensory device is its ability to translate the change it is to measure ... into some form of energy which can be amplified and used to drive the restoring cause ... The next part is what may be called the computing device and controller, which determines the amount and kind of energy to be released from the effector unit to restore equilibrium ... The final part is the power unit or effector (equivalent to the muscles in men and animals) which restores the state of equilibrium."

and computing machines as powerful metaphors for the functions performed by nervous systems as well as a basis for investigating human perception, adaptation and the acquisition of skill. Craik drew parallels between the operations performed by minds and machines, and suggested that perception and performance are based on mental models of the environment. The machine metaphor has proved to be particularly attractive to experimental psychologists. Craik was only able to enlist relatively simple machines, but his insight lies at the heart of the cognitive revolution that was to sweep through psychology. He worked with analogue devices as the digital computer was still embryonic in the 1940s.

Perhaps the most significant concept underlying these developments is that of information processing, which refers to the logical operations performed in transforming input data into outputs. Such operations can be described and analysed independently of the particular physical device that carries them out. This was made clear by David Marr (Figure 1.8) whose work brought together knowledge in computer science, psychology and physiology. Marr emphasised the importance of identifying the levels at which any information processing system, including the visual system, needs to be understood: "In a similar vein, trying to understand perception by studying only neurons is like trying to understand bird flight by studying only feathers: It just cannot be done. In order to understand

Figure 1.8 **2½-D sketch** David Marr (1945–1980) is represented by pattern tokens signifying the contours of his face, and his portrait is embedded in a representation of a cube; there are slight pattern differences to denote the boundaries between the three faces of the cube.

bird flight, we have to understand aerodynamics; only then do the structure of feathers and the different shapes of birds' wings make sense." At the top level is the description of computational theory. This involves stating the purpose and goal of the process, why it is appropriate in the context of other functions, and the general logic of the strategy needed to carry it out. Below this is the algorithmic level of representation. This level requires statements about the actual sequence of processes that take place, so that the computation is achieved. As an example, consider the process of multiplication. If you wish to multiply two numbers, then the goal of your computation is to generate a third number which is the product of the two original numbers. To achieve this, various logical procedures, called algorithms, could be employed. You could repetitively add one number to itself, as often as specified by the second number. Alternatively, you could make use of remembered information about the results of such additions, in the form of multiplication tables. Either procedure would in principle produce the correct answer. There will generally be several algorithms for any given computation. That is, the computational theory doesn't specify a particular means of achieving a goal, which must be determined independently. Marr referred to the first level as the 'what' and 'why' of the computation, and to the second as the 'how' of it. Independent of both computational theory and representation is the level of hardware implementation. Just as a given computational function can be carried out by a number of algorithms, so a given algorithm can be instantiated by various alternative physical devices. An algorithm for multiplication can be implemented in an electronic computer, a mechanical calculator, a brain or even an abacus. In terms of visual perception, Marr's analysis suggests that we should be careful not

to confuse the different levels of explanation. In addition, the starting point must be a proper description at the level of computational theory (such as aerodynamics in the example of birds' feathers cited above), because only this can provide a framework for understanding the operations at lower levels. If we lack a good general description of the purpose of a perceptual process, we are unlikely to understand its underlying logic, or the physiological hardware in which it takes place.

Computer architectures differ from the human visual system in a number of important ways. Visual systems rely on an extremely large number of relatively simple computations, occurring in parallel. The neural elements are multiply interconnected in a highly complex manner, which can alter as the result of experience. Since nerve impulses travel slowly compared with electronic signals, and nerve cells take time to recover from activity, biological computation is also slow. A typical electronic computer uses primarily serial processing, in which only one computation at a time can be performed, but at very high speed. There are rapid developments both in parallel computing and in the organisation of processes in the form of connected networks, but the scale of these is at present barely equivalent to even quite simple organisms. No doubt technological advances will in future enable the creation of computing systems that can in principle perform operations as complex as those underlying perception, in real time. However, it is possible that we simply do not yet have a clear enough idea about the way in which perceptual processes operate to be able to recreate them in a computer. We know broadly what the output of such processes should be (Marr's level of computational theory), but the level of algorithmic representation remains elusive.

The computer provides a useful metaphor for the brain because both can perform many different functions with the same structures. Despite these similarities we have to be careful not to take the metaphor too far. It has in fact proved very difficult to create anything resembling perception in a machine, despite the advances of recent years. Computer-based machine vision is a goal of many research groups, not least because of the considerable practical and economic advantages that it would give to industry and commerce. However, it has proved exceptionally difficult to achieve, despite a world-wide effort over many decades. In terms of crossing a road, the output of a light-sensitive device would have to be used to find out, amongst other things, how wide the road is, if there are vehicles and how long they will take to arrive. While solutions to some aspects of these problems have been achieved, an effective machine vision system would have to solve them all extremely rapidly and with a very high probability of success. It is certainly true that neither we, nor many animals, are wholly expert at road-crossing, as accident statistics sadly demonstrate. However, we do not know to what extent this reflects a failure of perception, rather than cognition; that is, the individual may see the environment correctly, but may make an inappropriate decision about the riskiness of an action due to attention or memory failure or errors of judgement. Nevertheless most people would still be happier to trust their safety to a trained guide dog rather than to the most advanced of machine vision systems.

Reference notes

In the reference notes we will provide guidance to other books and articles that address issues we raise. We will also provide the sources of quotations given in the text, so that interested readers are able to follow up aspects of our arguments by referring to the original sources. A full list of references is given in the bibliography. In addition to the books cited here there are many resources available on the internet. One of the aspects of this book, which is not shared by other introductory textbooks on perception, is its emphasis on the ways in which our ideas about the senses have developed over time. This is evident from this chapter but it is amplified in the topics covered in the ones that follow. Indeed, much of the historical interest in vision has concerned its abnormal function rather than its normal, reliable delivery of the state of the external world.

General features of perception are covered in almost all introductory textbooks on psychology and there are books devoted to this topic. Mather's (2008) *Foundations of sensation and perception* and Goldstein's (2009a) *Sensation and perception* are good examples. Both are well illustrated and both focus on the physiological underpinnings of perception. The flavour of contemporary research on visual perception is provided in *Basic vision* by Snowden, Thompson and Troscianko (2012), in *Why we see what we do* by Purves and Lotto (2010), and in *Seeing* by Frisby and Stone (2010). Most texts on perception give greatest prominence to vision, as most of the research on perception has been directed to this sense. Goldstein has also edited an *Encyclopedia of perception* (2009b). Many of the issues introduced in this chapter engage several different disciplines. Theories of philosophy have drawn upon perception to varying extents, and philosophers have added to our understanding of it. The edited book by Symons and Calvo (*The Routledge companion to philosophy of psychology*, 2009) has a section devoted to perceptual experience in addition to a wide range of other influences like the historical background to psychology. Philosophical accounts of perception have tended to be concerned with epistemology – deriving evidence about knowledge of the world. Such an enterprise has concentrated on human perception. Because we wanted to avoid an overemphasis on language and human perception, we have considered how guide dogs can interact with their perceptual world and even assist in navigating humans through space. It also provides a vehicle for discussing species differences in perception. The visual capacities of dogs are provided in the article by Miller and Murphy (1995). Most investigations of eyes have been addressed to the image-forming variety of vertebrates but the full splendour of evolutionary adaptations for transducing light energy has been detailed in Land and Nilsson's (2001) book *Animal eyes*.

We are emphasising perception in a three-dimensional environment so that vision operates in concert with other senses. *Active vision* is the title of a book by Findlay and Gilchrist (2003), and it concentrates on the interplay between eye movements and vision. This aspect of active exploration of the world with the eyes is also the thesis presented in Land and Tatler's (2009) book *Looking and acting: Vision and eye movements during natural behaviour.* The example we

have chosen of the perceptual demands placed upon crossing a road deals with whole body motion as well as eye movements. Gibson (1966, 1979) stressed the importance of vision in guiding action and also that vision cannot be divorced from the dimension of time, and this approach has been continued by Cutting (1986) in his book *Perception with an eye for motion.*

The classical psychophysical methods were formulated by Weber and Fechner in the nineteenth century. Some of their works concerning measurement in perception have been translated into English. Weber's Latin and German texts on touch are now available (Ross and Murray, 1996) and the first volume of Fechner's (1860/1966) two-volume *Elements of psychophysics* is available in English. The 150th anniversary of the publication of Fechner's *Psychophysik* has been marked by the essays in Solomon (2011). Good accounts of contemporary psychophysics can be found in Sekuler and Blake's (2005) *Perception,* which also covers psychophysical scaling. Stevens (1975) presented an account of his scaling methods. He had earlier presented his alternative to Fechner's law in 'To honor Fechner and repeal his law' (Stevens, 1961). A modern attempt to reconcile the two approaches to psychophysics can be found in Billock and Tsou (2011).

Physiological indices of perception have a long past and a vibrant present. One of the oldest techniques is EEG, which is still widely applied both clinically and for research purposes. These measures, together with descriptions of modern brain imaging techniques, can be read about in *Neuroscience: exploring the brain* by Bear, Connors and Paradiso (2006). The insights that have been derived from fMRI are displayed in *Human brain function* edited by Frackowiak *et al.* (2004). Walsh and Pascual-Leone (2003) describe the applications of TMS to a range of psychological tasks in their book *Transcranial magnetic stimulation: a neurochronometrics of mind*. Some remarkable images derived from new techniques of staining and imaging neural tissue can be seen in Schoonover's (2010) *Portraits of the mind. Visualizing the brain from antiquity to the 21st century*.

The text with which Craik's portrait is combined is from his book *The nature of explanation* (1943). A selection of his papers can be found in *The nature of psychology* (Craik, 1966). Marr's book *Vision* was published posthumously in 1982, and it continues to have an influence on the ways vision is thought about. Marr appreciated that understanding vision can only be achieved when the purpose that it serves is understood. This was considered to require integrating studies from physiology, psychophysics and artificial intelligence. Gordon's (2004) *Theories of visual perception* is one of the very few books concerned primarily with theories rather than experiments and it provides an excellent introduction to the problems that any theory must address (see also Wade, 2012a).

The 'perceptual portraits' of Gibson, Weber, Fechner, Stevens, Craik and Marr are from *Psychologists in word and image* (Wade, 1995), and many more can be found at http://neuroportraits.eu/ .

2 Fundamentals of vision

Modern accounts of vision draw on the past in innumerable and often unstated ways. The distinctions we draw between the physics of light, physiological responses to light and the psychology of perception, represent the outcomes of hard-won battles. The theories applied to each of these remain matters of investigation and debate. We are in a better position to appreciate the massive advances that have been made in understanding the nature of light, image formation, visual anatomy and physiology, and visual phenomena themselves when we can place them in the context of the past. The historical perspective is often overlooked or neglected in books on perception, which is regrettable because it implies that we now have a specially privileged viewpoint, superior to those of the past. In fact, the same theoretical issues often recur, disguised by new jargon to appear different. A historical perspective serves to emphasise the continuing evolution of ideas, concepts and theories in the light of critical evaluation and experimental evidence.

The importance of appreciating the heritage of vision was summarised eloquently by Johann Wolfgang Goethe (Figure 2.1) 200 years ago: "We cannot clearly be aware of what we possess till we have the means of knowing what others possessed before us". Goethe was not only a poet and playwright of enormous popularity but he also wrote on philosophy, science and the senses. In line with many Romantic philosophers, he rejected the experimental approach to the study of nature because it was considered to be too constrained. In its place he proposed the astute and intuitive observation of natural phenomena, setting in train the method of phenomenology. This is best seen in his *Theory of colours*, which contrasted his observational approach to colour with what he considered to be the physicalism of Newton. Goethe chose to observe and describe instead of experiment on colour vision. He distinguished between what he called physiological colours (the experience of colour) and physical colours produced by optical refraction. Goethe's theory of colour was never taken seriously by the scientific community, but his observations have rarely been challenged. He reported many novel phenomena, like positive and negative colour afterimages, irradiation, colour shadows, and colour blindness, in addition to contrast effects – both in the chromatic and achromatic domains.

Goethe was one of many students of vision who have added to our understanding of its processes, but he was by no means the first. Writers in classical Greece

Figure 2.1 **Romantic eye** The silhouette of Johann Wolfgang von Goethe (1749–1832) is composed of a pattern of contrasting black and white, diamond shapes which are rather like those in a figure from his *Theory of colours*. The sizes of the elements vary very slightly, as can be seen by a good observer, to reveal Goethe's acute left eye. The pattern, if observed for some time, might even appear in pale colours; these are called subjective colours and they were also investigated by Goethe.

discussed and debated the nature of light and sight and some of their ideas will be described in this and the next chapter. However, the dawning of the experimental approach to vision had to await an understanding of how images are formed in the eye and the anatomy of the eye itself. These developments occurred at the beginning of the seventeenth century, and rapid advances were made thereafter. This is the starting point that Edwin Boring (Figure 2.2) adopted in his book *Sensation and perception in the history of experimental psychology*. Boring was aware that establishing an origin for the study of the senses was problematical since the history of perception is as long as that of description itself. Nonetheless, his spiral of history lists the major figures from 1600 to 1900. An implicit feature of the spiral is that one sense – vision – dominated all the others. A preoccupation with the sense of sight was evident throughout the period surveyed by Boring and it remains so to this day. Many of the people placed on the spiral will appear in the pages that follow, and the spiral itself provides a useful indication of the periods in which they were active.

Of course, studies of vision did not begin at the centre of the spiral (around 1600) because an observational tradition was well established before that. Writers have always remarked on their own visual experiences, so that a large body of descriptions of observations has accumulated. Science involves recording and interpreting natural phenomena. In contemporary science, the records are the results of experiments and the many and varied phenomena are posited in well-defined compartments, like physics, physiology, and psychology. These compartments are a relatively recent convention, as are the specialists who labour under titles such

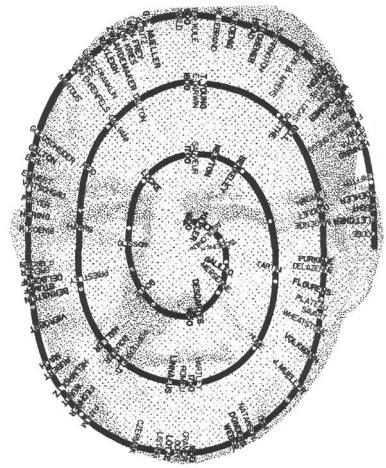


Figure 2.2 **Historian of perception** Edwin Garrigues Boring (1886–1968) is represented in the spiral of history, with its significant figures, as displayed in the inner cover diagrams that sandwiched his book *Sensation and perception*. As with all good historians, he barely intrudes on the history he presents, but his presence is evident throughout. The spiral starts at 1600 and Kepler is the first name shown; many of those who will be encountered in the following pages are listed thereafter, like Scheiner, Descartes, Locke, Newton, Berkeley, Reid, Dalton, Young, Purkinje, Wheatstone, Helmholtz, Mach, Wundt, Hering, James, Wertheimer, Koffka, Köhler and ending at 1900 with Dodge.

as physicists, physiologists, and psychologists. Neither the phenomena nor the practitioners were so clearly defined in the distant past. In antiquity science, if such it should be called, was based upon describing and classifying observations of naturally occurring events. The sense of science was sight, and vision itself was an integral part of the development of science. What was seen could be described, classified, and even subjected to mathematical analysis. The onset of Boring's spiral also signals a graphical transition as those who lived before the seventeenth century will be represented with conventional portraits and those after that date with perceptual portraits. This strategy has an additional value because we rarely have portrayals of the ancients (like the Greek students of vision) for which there is any evidence of their likenesses to the persons portrayed. Thus, in the main, the conventional portraits are based more on imagination than observation.

Recording the consequences of perception started long before written reports were kept, and so art could be encompassed within this purview. In addition to the historical perspective, considering perception from the point of view of artistic representation can provide a valuable counterpoint to current scientific approaches, especially since they share much common history. In this chapter, we also review the changes that take place during perceptual development and when some failure occurs. When we are born we do not see the world as sharply as when we are adults, and a host of other changes take place during development. Moreover, we do not all see the world in the same way, either as a consequence

of structural differences (like the optics of the eyes) or the ways in which signals from the eyes are processed in the brain. Thus, we consider the nature of visual perception from a diverse range of perspectives and we emphasise the importance of knowledge gained from a variety of different sources. Many of the current issues in visual perception have ancient roots. Appreciation of these is valuable not only to avoid past errors, but also to provide a basis for asking new questions.

Visual perception and the physical environment

The investigation of perception necessarily includes the definition of those characteristics of the physical environment that a perceiver is capable of detecting. We need to know at the least what sorts of physical event can act as an effective stimulus for perception and that a given event is capable of being perceived, before engaging in discussion of its behavioural significance. For example, before expending effort in training a guide dog to respond to an environmental feature such as colour, it would obviously be necessary to establish that the dog could detect the relevant feature in the first place. The point may seem obvious, but it is unwise simply to assume that other people, let alone other species, perceive what you do.

Perception exists as one means of ensuring survival. If our actions are not guided by reliable information about the world (within the limits of sensory detection), then we may make errors and put ourselves and others at risk. The important questions for a psychologist concern how we come to gain information about the environment, the way in which such information is represented in perception and the limits to our ability to perceive correctly. The ability to ask such questions rests on an understanding of how certain wavelengths of electro-magnetic radiation (light) come to initiate the subjective experience of seeing. This understanding was arrived at only after many centuries of enquiry and speculation.

Light and sight

All sciences are based on observation. Naturally occurring or experimentally manipulated events are observed and if the observations prove to be consistent they become phenomena. That is, if different observers describe the events in much the same way then they are considered to be aspects of the world rather than aspects of the observer. The phenomena are accumulated and classified in order to interpret them and determine lawful relations between them. The history of vision is somewhat unusual because it is the history of observation itself. For most of its history, the phenomena of vision were those based on observation of natural events like rainbows and reflections, together with the consequences of diseases like cataract and failing sight in old age. Thus, investigations of both the physics of light and the physiology of sight added to the range of visual phenomena, although these two areas were not initially distinguished. The psychological dimension was gradually integrated with these. Vision became an experimental discipline rather late. For most of its long history it has drawn upon observations

Figure 2.3 **Optics, catoptrics and dioptrics** Hero of Alexandria (c. 10–70 AD) is shown in direct vision (optics), in reflection (catoptrics) and with his face partially submerged in water so that the lower part is refracted (dioptrics).

of naturally occurring events, with relatively little in the way of manipulating the conditions under which they occurred.

Optics is now considered to be the branch of physics dealing with the properties of light – how it is propagated through different media, how it reflects from surfaces, and how it is refracted at the boundaries between media. These properties have been studied for well over two thousand years, and the terms to describe them – optics, catoptrics (reflections), and dioptrics (refractions), that were introduced by Hero of Alexandria in the first century AD – are still in use (Figure 2.3). A distinction is now drawn between physical optics and vision – between light and sight – but this is relatively recent. It is only since the seventeenth century, when the optical functions of the eye were correctly described and when its gross anatomy was accurately depicted that the two have been distinguished. This is why Boring's spiral of history (Figure 2.2) started at around 1600.

Among the many philosophers and mathematicians in ancient Greece, the most penetrating analysis of optics that has survived was written by Euclid (Figure 2.4) around 300 BC. Euclid's optics were both physical and psychological; the rays

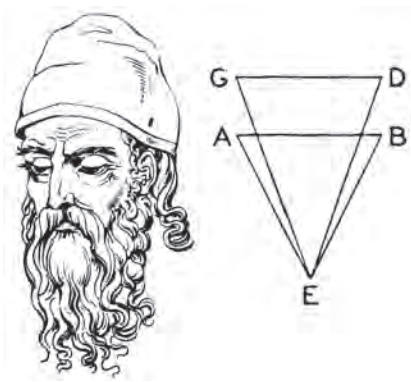


Figure 2.4 Euclid (c. 323–283 BC) together with a diagram of lines of light from the eye (E) to two objects of the same size at different distances. Euclid equated perceived dimensions to the angles subtended at the eye.

Figure 2.5 Left, Ptolemy (c. 100–170) and right, Ibn al-Haytham (c. 965–1039) provided the most detailed accounts of optics and its association with vision between the time of Euclid to the beginning of the scientific revolution in the seventeenth century.

in his diagrams represented not only light but also sight. For example, Euclid's book entitled *Optics* is concerned with describing visual space in terms of visual angles, so that the perception of space was analysed geometrically. For example, the visual cones shown in Figure 2.4 represented both the optical projection and the perception of two equal extents; the visual angle of the nearer object (AB) is greater than that of the distant one (GD) and Euclid contended that AB was seen as larger than GD. The two most important treatments of optics between Euclid and the early seventeenth century were first by Ptolemy and then by Ibn al-Haytham, who is more widely known by his Latinised name of Alhazen (Figure 2.5). The *Optics* of Ptolemy is both experimental and psychological, dealing with colour as well as space.

Ptolemy appreciated that light rays should be thought of as continuous rather than discrete. His analysis of visual size was radically different from that of Euclid, as it was based on psychology rather than physics. Ptolemy argued that visual size cannot be equated with visual angle, and introduced the concept that perceived size was derived from visual angle and distance; that is, he addressed the issue of perceptual constancy – that objects appear to have about the same dimensions despite the projected dimensions at the eye. He was in agreement with Euclid about the variations in visual acuity throughout the visual pyramid (rather than visual cone). The two pyramids of vision (one for each eye) needed to be integrated and he conducted experiments with a board in order to study this binocular combination. Ptolemy also realised that errors do sometimes occur in vision, and he was one of the first writers to provide a detailed account of illusions. Indeed, he devoted over one third of Book II of his *Optics* to errors of sight; they were classified, and then considered under the headings of colour, position, size, shape, and movement. Ptolemy initiated a reconciliation between physical and psychological analyses of vision which was further amplified by Ibn al-Haytham.

Ibn al-Haytham was a mathematician who hailed from Basra and later worked in Cairo. He was familiar with the writings of Euclid and Ptolemy and Ptolemy's

influence was particularly marked. Ibn al-Haytham adopted a reception theory of light, conducted experiments on image formation in a pinhole camera, extended Ptolemy's experiments on binocular vision, and classified a wide range of visual phenomena. In short, he integrated the three theoretical strands in optics that had until then been dominant but distinct – Aristotle's observations and philosophical stance, Euclid's mathematical treatment in terms of ray diagrams, and the physiological approach of Galen. Ibn al-Haytham's book on optics was translated from Arabic into Latin in the thirteenth and again in the sixteenth century.

Early ideas about light

We now accept that light is emitted by incandescent sources like the sun, and reflected from objects to enter the eye, so initiating the process of seeing. However, for millennia it was thought that light was emitted from the eye to make contact with objects and returned to the eye with their images; vision itself was thought to occur in the lens. After all, is it not the case that we cease to see when we close our eyes, thus preventing the emission of light from the eye? Moreover, what we are aware of depends upon where we direct our eyes. Most Greek speculations about sight and the senses generally incorporated elemental philosophy so that fire, earth, water, and air permeated perception. Touch was often taken as the most important sense, and the one relative to which others could be related; qualities associated with it, like hot, cold, moist, and dry were thought to be common to all the senses, and were in turn linked to the four elements. Thus, vision was generally considered to involve some process of contact between the eye and objects and several means of achieving this contact were advanced. These included various versions of emission or extramission theories, in which light originated in the eye and was projected from it. Reception or intromission theories, in which light travelled from objects to the eye, were also advanced, as were speculations incorporating aspects of both emission and reception. Emission theories could have been founded on the experience of light when pressure is applied to the eye and they are consistent with the cessation of sight when the eyes are closed.

In around 500 BC, Alcmaeon (Figure 2.6) described seeing flashes of light when the eye was pressed or received a blow and he created the enduring metaphor of a light being a 'fire in the eye'. He wrote: "the eye obviously has fire within, for when one is struck [this fire] flashes out. Vision is due to the gleaming – that is to say, the transparent – character of that which [in the eye] reflects the object". This phenomenon proved very difficult to encompass within reception theories of light, and is largely responsible for the longevity of emission theories. Whereas it could be demonstrated that light is not seen when the eyes are open in darkness, the experience of light flashes with pressure to the eyes was an enduring problem for theories of light that required reception by the eye. A theory based on light passing from objects to the eye was advanced by Democritus (Figure 2.6). It could account readily for the absence of sight with eye closure but not for the experience of light when external pressure was applied to the eye. What was received by the eye was often more than light, but some image of the object itself. According

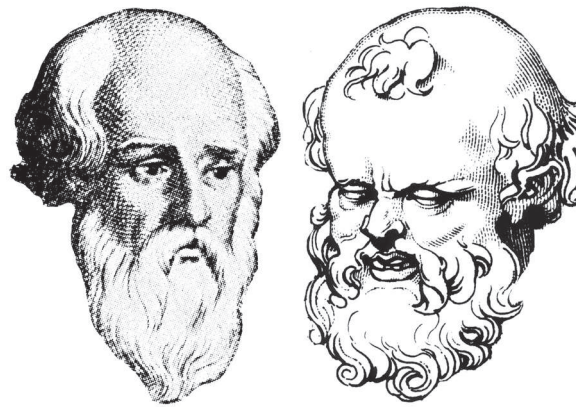


Figure 2.6 Left, Alcmaeon (c. 500 BC) and right, Democritus (c. 460–370 BC).

to Democritus all nature was composed of atoms in motion, and images were continually emitted from objects to compress the air and carry impressions to the eye. The impression was like a copy of the object that could be received by the eye. The images carried with them the constant features of the objects, and for Democritus these included their three-dimensionality.

A combination of emission and reception was proposed by Plato (Figure 2.7), although his theory of vision was always subservient to his philosophy of ideal forms. Plato suggested that light was emitted from both the eye and objects, and vision took place externally where these two streams united. He suggested that we cannot see at night because light is extinguished in darkness, just as heat and dryness are extinguished by cold and dampness. Aristotle (Figure 2.7) was scornful of this speculation, stating that neither heat nor dryness were attributes of light. For Plato, as for Aristotle, it was not light but colour that was the principal source of interest in vision. Plato distinguished between light and colour, considering

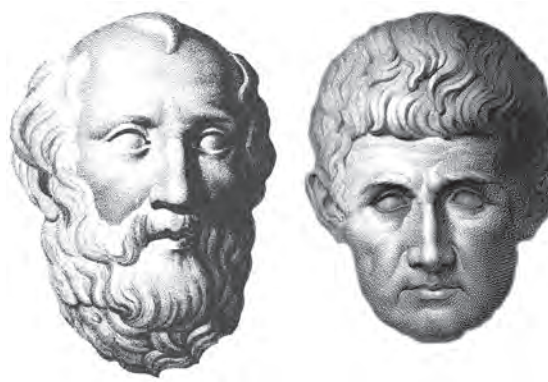


Figure 2.7 Left, Plato (427–347 BC) and right, Aristotle (c. 384–322 BC).

that light had its ultimate source in the sun, but colour was a property of objects themselves.

Aristotle's theory is more in line with modern conceptions of light. His interests were in observation, and the phenomena he experienced directed the interpretations he proposed. Thus, he queried the emission theories by the simple expedient of testing a prediction that would follow from them: if light was emitted from the eye then vision should be possible at any time the eyes were open, including night time. The fact that the prediction was not supported led him to suggest an alternative theory of the nature of light. Similarly, he distinguished between vision and touch by noting that an object in physical contact with the eye could not be seen. His alternative interpretation was that vision is the result of some movement in the medium separating the eye from the objects perceived. Aristotle's reception theory was not widely accepted, in part because it could not account for the experience of light in darkness when pressure was applied to the eye.

Euclid adopted an emission theory of light and vision. Vision was restricted to the cone of rays emanating from the eye and meeting the objects within it (see Figure 2.4). The geometrical projections to these objects were lawful, and this lawfulness was applied to vision, too. Thus, Euclid provided not only an account of optical transmission through space, but also a geometrical theory of space perception itself. The perceived dimensions of objects corresponded precisely to the angles they subtended at the eye, and illumination of those objects had its source in the eye. The emissions from Euclid's eyes were referred to as visual rays, and their properties were conflated with a number of phenomenal features. The visual rays were thought to be discrete, and so small objects could fall between them, and remain unseen; that is, there is a limit to the dimensions of objects that can be detected, namely, a threshold for visual acuity. Moreover, those objects illuminated by rays in the centre of the visual cone will be seen more clearly than those towards the edge; that is, direct (foveal) vision has better acuity than indirect (peripheral) vision. It followed, as Euclid stated, that nothing could be seen at once in its entirety, implying that the visual rays would move over an object (by moving the eyes) in order to see all its features.

The belief that light was emitted from the eye rather than transmitted to it was held for over 2000 years until, in 1604, Johannes Kepler (Figure 2.8) described how light passed through the eye to form an image on the retina. The similarities between image formation in a simple camera obscura (literally a dark chamber, like a pinhole camera) and the eye were remarked upon in that period, and the emission theory of vision was replaced by a reception theory. Indeed, at that time it was considered by some that the problem of vision had been solved by the appreciation of the image-forming properties of the eye: the picture in the eye was like the scene imaged, and so it corresponded to perception. Kepler was more circumspect in treating the relationship between the inverted and reversed retinal image and perception: "I leave it to the natural philosophers to discuss the way in which this image or picture is put together by the spiritual principles of vision". Natural philosophers have not subsequently spoken with a single voice

Figure 2.8 **Kepler's retinal image** Johannes Kepler (1571–1630) described the optical image forming characteristics of the eye, introducing the concept of the retinal image. He is shown in a diagram of the microscopic structure of the retina which is in turn enclosed within the words RETINAL IMAGE in the upper half. However, Kepler showed that the retinal image is inverted and reversed with respect to external objects and so these words are also inverted and reversed in the lower half of the illustration.

on this problem, but they have appreciated that physical optics is not the solution to vision.

In addition to the description of image formation in the eye, Kepler made many other contributions to optics, both experimentally and theoretically. Amongst them was the formulation of the basic principle of photometry that the intensity of light diminishes with the square of the distance from the source. So if a source has a luminance of x at 1 m, then the measured intensity will be $x/6^2$ at 6 m, $x/10^2$ at 10 m, and so on. This 'inverse square law' applies not only to light but also to gravity, for which it was first demonstrated by Newton. In addition to his book on optics of 1604, Kepler published another on *Dioptrice* (Dioptrics) in 1611; in this he devoted considerable attention to refraction, but he did not determine the general sine law. Willebrord Snell (1580–1626), in an unpublished manuscript written around 1621, described the relationship between angles of incidence and refraction (the ratio of the sines of the angles of incidence and refraction is a constant – the refractive index – for any given material through which light passes). The subsequent technical advances in optical instrument manufacture were based on this relationship, which became known as Snell's law.

Snell's law was elaborated by René Descartes (Figure 2.9) in his *Dioptrique* published in 1637. Light, according to Descartes, acted like a mechanical force which is transmitted through transparent media. His theory of light attracted much criticism in his day because of the inconsistencies it embraced. On the one hand he argued that light was propagated instantaneously, and on the other that it varied its velocity according to the density of the medium through which it travelled. Nonetheless, Descartes did much to integrate optical image formation,

Figure 2.9 **Descartes' method** René Descartes (1596–1650) presented his analysis of optics in his book on the scientific method, the title page of which is shown with his portrait.

the anatomy of the eye and a speculative theory of vision, as we will see later in this book.

Until the seventeenth century the phenomena of light were restricted to the classical divisions of optics, catoptrics and dioptrics described by Hero of Alexandria (see Figure 2.3). These were added to by Franciscus Maria Grimaldi (Figure 2.10) who described diffraction of light. Diffraction occurs when light passes sharp boundaries and appears to bend. Grimaldi demonstrated the phenomenon by partially blocking sunlight passing through two small apertures: bands of colour could be seen in the shadow area. On the basis of diffraction he suggested that light might act like a liquid, flowing in waves. Robert Hooke (1635–1703), independently of Grimaldi and at about the same time (1675), examined diffraction produced by placing a razor's edge in the path of a narrow beam of sunlight. Hooke described the bands that were visible in the shadow region but he did not remark on any colours seen in them. A decade earlier, Hooke had noted the coloured rings in thin transparent films and interpreted them in terms of vibrating pulses of light. On the basis of his observations and deliberations he listed three properties of the pulses or waves of light: "First, I found it ought to be exceeding *quick* … Next it must be a *Vibrative motion* … Thirdly, That it is a very *short vibrating motion* … And these, I think are the three principal properties of a motion, requisite to produce the effect call'd Light in the Object". He also speculated that the vibrations were transverse to the direction of propagation, and that the waves were like those generated by a stone dropped in water. Hooke's concept of light was elaborated in his lectures delivered at the Royal Society, and his diagram representing the propagation of light is shown in Figure 2.10.

Wave theory was supported and extended by Christiaan Huygens (Figure 2.11): he proposed and illustrated the wavefronts that could be produced by points on luminous sources, and he made an analogy between light and sound; diffraction was analysed in terms of the wavefronts originating at the aperture. Since the time

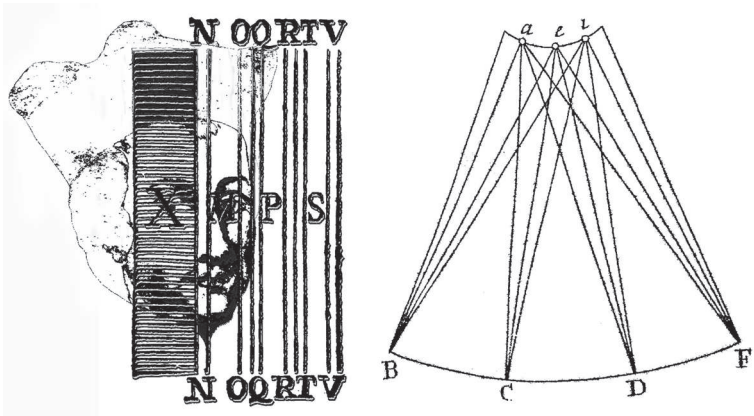


Figure 2.10 Left, **Grimaldi's diffraction** Franciscus Maria Grimaldi (1613–1663) is shown together with the diagram from his book on light representing diffraction bands. Right, Hooke's diagram of the propagation of light by analogy with a series of syringes acting on the base of a cone: "Suppose, a, e and i, were the three Holes of these Syringes, and B, C, D, F, were the bottom of the Cone; if the Syringe e were moved, it would propagate a Motion, or force to every point B, C, D, F, of the bottom B, F, so likewise if the Syringe a, or the Syringe i, were moved singly, each of them would propagate a Motion to every one of those Points B, C, D, F, of the Base B, F, and to every other Point of it. And consequently, all those Syringes being moved together, must every of them influence or move every Point of the Base B, F, with its own distinct Influence. Now the same thing that will happen in this Case, in the pressure or Motion of the Water against the bottom, will also be in the Case of Light; for if we suppose a, e, i, to represent the Body of the Sun, and B, C, D, F, the Surface of a Body inlightned by it; It is most certain, that every Point of the Base or Object B, C, D, F, is influenced by every Point of the Superficies of the Sun that shines".

Figure 2.11 **Wave theorist** Christiaan Huygens (1629–1695) is portrayed twice. On the left, he is shown in his illustration of the waves that could emanate from a candle flame, taken from his book on light, the title page of which is also presented. On the right, he can just be seen in an abstract pattern of waves.

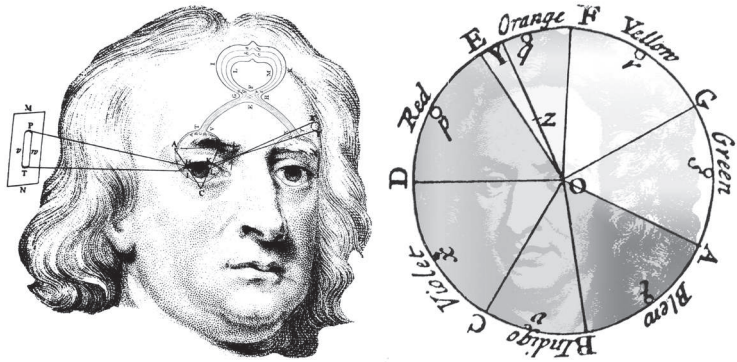


Figure 2.12 **Spectral visionary** Isaac Newton (1642–1727) analysed the spectrum produced by passing a narrow beam of sunlight through a prism and projecting it onto a screen in an otherwise dark room. On the left, his illustration from the Opticks is superimposed on the portrait. His speculations regarding the pathways from the eyes to the brain are also shown. On the right he is enclosed within his diagram of the colour circle. The colours were arranged so that opposite pairs when mixed together would give a colourless combination.

of Grimaldi, theories of light have been informed by analogies with sound. In the eighteenth and nineteenth centuries, the alternative view that light consisted of small corpuscles was challenged by phenomena that conformed more closely to the action of waves rather than particles. Sound was analysed in terms of waves and a similar approach was suggested for light.

Isaac Newton (Figure 2.12) proposed that light consisted of small corpuscles which collided with one another. Thereafter, the theoretical contrast was between Huygen's wave theory and Newton's corpuscular theory of light. The latter was dominant throughout the eighteenth century, so much so that it became dogma, and those who challenged it were roundly condemned for doubting the genius of Newton. However, it was the nature of colour rather than that of light itself that was illuminated by Newton's thought. Euclid's *Optics* was concerned with the geometry of space and did not mention colour. Colour compounds the physical and psychological dimensions of optics, although an appreciation of this was only possible when the physical properties of light had been elucidated. In his *Opticks* (1704) Newton illustrated and interpreted the spectrum produced by sunlight passing through a prism: differences in colour were related to differences in the amount they were refracted by the prism. He also noted that colour was a subjective experience rather than a property of light: "For the Rays to speak properly are not coloured. In them there is nothing else than a certain power and disposition to stir up a sensation of this or that Colour".

That power was dependent in part on the degree to which sunlight was refracted when passing through a prism. A wide range of colour phenomena was available to Newton before he conducted his prismatic experiments, including the prismatic colours themselves. The phenomena concerned colour mixing, afterimages, colour contrasts and colour shadows; they had accumulated over centuries and

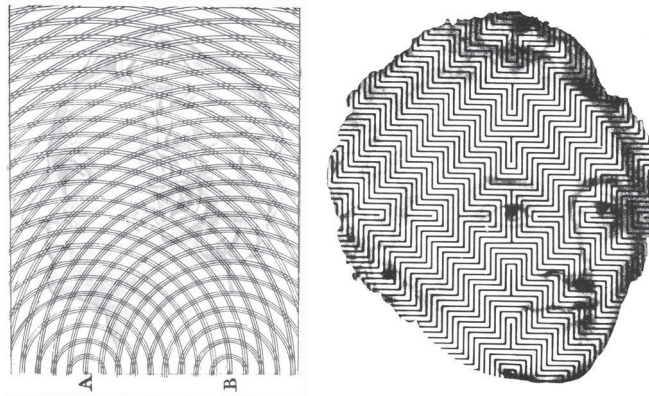


Figure 2.13 **Interference and astigmatism** Left, Thomas Young (1773–1829) can be seen in his representation of interference: light passing through two small apertures (A and B) interacts producing peaks and troughs rather like waves of fluid. Right, his portrait is contained within a pattern of vertical and horizontal contours which will differ in clarity if you have astigmatism.

spawned competing theories. Most theories were based on the Aristotelian view that colour was a property of bodies, and that it was carried by light through the transparent medium to the eye.

Newton was to provide a unified theory of light and colours: white light was not unitary but compounded, in precise proportions, of many components that differed in their refraction. He was able to analyse the prismatic colours with an armoury of subtle experimental procedures. For example, small sections of the prismatic spectrum could be isolated and mixed with other parts, so that the rules of their mixture could be determined. He wished to avoid the speculations entered into by his rival, Hooke, and to base his theory on experiment rather than hypothesis. Indeed, Newton opened his *Opticks* thus: "My Design in this Book is not to explain the Properties of Light by Hypotheses, but to propose and prove them by Reason and Experiments". From this time onwards the aim of the scientist was to proceed by observation and experiment.

By the beginning of the nineteenth century, corpuscular theory was accepted as a conceptual convenience, but one which found increasing difficulty in accounting for phenomena like diffraction. Thomas Young (Figure 2.13) had been fascinated by the nature of sound and light since his student days. In addition to applying wave theory to account for diffraction, he introduced a new optical phenomenon, interference. He described it thus: "When two Undulations, from different Origins, coincide either perfectly or very nearly in Direction, their joint effect is a combination of the Motions belonging to each". Young's illustration of the phenomenon is shown in Figure 2.13. Young was also fascinated by visual optics and he measured the dimensions of his own eyes and charted the ranges over which they could focus sharply (a process called accommodation). Descartes had earlier speculated that accommodation occurred as a consequence

of variations in the curvature of the crystalline lens, but the alternative of corneal curvature changes was still entertained. In 1801, Young demonstrated that the former interpretation was correct by noting that accommodation persisted when the eye was immersed in water, eliminating corneal refraction (at the boundary between the eye and the external medium). In addition, patients without a lens were unable to accommodate to objects at different distances. He also made detailed measurements of the dimensions of his eyes using an optometer of his own devising. Young described the optical aberration of astigmatism, the condition in which lines in different orientations cannot be brought to a focus in the same plane; when one set of lines is sharply focused the other appears blurred, and vice versa. Astigmatism is usually a consequence of the curvatures of optical surfaces in the eye not being parts of spheres. Most commonly it is differences in the curvature of the cornea in different orientations that leads to astigmatism, but there is another form that is due to slight changes in the accommodative state of the lens.

With the appreciation that light could be considered as a physical property, and that its reflections and refractions followed physical principles, its study became the province of physicists, whereas the examination of sight was pursued by physiologists and philosophers. The separation of the physics of light from the philosophy of sight was to reflect the ancient schism between materialists and idealists: light was an external, material phenomenon whereas sight was internal and subjective.

Art and representation

Prior to laboratory studies of vision, the practitioners of this sense were artists. They manipulated colour and contour in ways that are not dissimilar to the approaches adopted by modern scientists. Indeed, the first systematic attempts at mixing pigments to produce new colour experiences were made by artists. Pigments have been employed to represent objects for tens of thousands of years and rules for the combinations of colours have a shorter but still considerable history. With this legacy it is surprising that the science of colour did not materially influence the practice of art until the nineteenth century. The situation is otherwise for spatial representations, where issues of contour interactions were addressed by artists before they were appreciated by scientists. Considering this history, it is surprising that phenomena disclosed to the scientist's eye are often reported as discoveries. In many cases, those practitioners of vision, artists, have discovered and manipulated the same phenomena often centuries before they came under scientific scrutiny, and often with greater subtlety.

Paradoxically, pictorial stimuli were latecomers to the discipline of visual science. Prior to the use of a variety of instruments for presenting stimuli to the eyes, many of which were invented in the early nineteenth century, vision was examined with respect to objects rather than representations of them. That is, the scientific study of vision was concerned with solid objects whereas visual artists were constrained by the pictorial surface. The instrumental revolution

transformed not only our vision of pictures but also our picture of vision. Pictures could be paired in a stereoscope, to appear in depth, or presented in rapid succession, to appear in motion, thereby replacing the two dimensions missing in static and single pictures – depth and motion. Theories of spatial vision developed thereafter emphasised its constructive nature, and placed greater emphasis on the manipulations that could be made of two-dimensional stimuli. One consequence of this shift from solidity to surface was a neglect of the centuries of surface (pictorial) manipulations made by artists. Thus, a large body of spatial phenomena was rediscovered in the confines of the visual laboratory. The perception and representation of motion, on the other hand, had a different history. Motion had been frozen in pictures and few novel techniques of alluding to it were developed. Prior to the invention of instruments for generating apparent motion, moving objects had been the source of scientific study. Thereafter, the synthesis of motion from sequences of still images was to have profound effects on popular culture as well as on artistic representation.

We will examine the development of understanding colour first and then consider the ways in which space has been represented in pictures.

Colour

Within Greek science, colour was considered in terms of an analogy with the four elements (earth, air, fire and water), and of these fire and water were dominant. The numeration of four basic colours was clearly stated in the fourth century BC by Democritus, and these were white, black, red, and green; other colours were derived from mixtures of these. Moreover, he adapted the concept of pores in the eye to account for colour vision: only when the geometrical shapes associated with particular colours corresponded to those of the pores would colour be experienced. Both Plato and Aristotle considered that colour was of paramount importance in perception, and that it could be dissociated from light. They appreciated that pigments could be extracted from certain substances, and they were well aware of the ways in which they could be mixed by artists. However, they stressed the importance of black and white: Plato treated them as opposites, and Aristotle considered that all colours could be made up from these two. The observation that no colours appeared as light as white or as dark as black might have led to this speculation. Consequently, despite the equation of colours with the four elements, black and white tended to dominate the analyses of colour until the time of Leonardo da Vinci.

The situation changed with investigations of prismatic colours. Although the occurrence of colours due to light passing through or reflected from glass or crystals had been known about since antiquity, their experimental examination had not been undertaken. A preliminary study of the prismatic spectrum was conducted by Thomas Harriot (c. 1560–1621) at the beginning of the seventeenth century, and it was subjected to more detailed analysis later in the century. Descartes treated the analysis of colours visible in the rainbow in his discourses on *Meteorology*. His mechanistic interpretation of visible colours was in harmony with his concept

of light generally: colours corresponded to different rates of rotation of bodies in the medium.

Newton demonstrated experimentally that sunlight (white light) is made up of rays that can be bent or refracted by different amounts when passing through a prism, so forming the visible spectrum (see Figure 2.12). He conducted experiments with a prism from 1666, reporting the results in a paper to a meeting of the Royal Society on 6 February 1672; it was published in the *Philosophical Transactions* later that year. The paper stimulated widespread interest as well as opposition, particularly from Hooke and Huygens, who argued that there were only two basic colours. The controversy stirred by the theory so perturbed Newton that he did not publish his full account of it until his *Opticks* appeared in 1704. By this time, he was President of the Royal Society. Newton was aware that the range of refractions from a prism was continuous and yet the colours seen were restricted in number. He reported seven colours – red, orange, yellow, green, blue, indigo, and violet – and he arranged them in a particular circular sequence, after the manner of the musical scale (Figure 2.12). Coloured lights could be combined in such a way that their compound could be defined with respect to the colour circle. Combinations of primary colours that were opposite one another on the circle produced a whitish compound that was positioned near the centre.

Much of the subsequent debate focused on the nature and number of primary colours. For Newton, who introduced the term in 1672, they were the discrete colours that could be seen in the prismatic spectrum, despite his appreciation of continuity across the spectrum. With the formulation of the colour circle they also became the colours from which compounds could be derived. Young proposed that red, green, and blue were primaries, and that the vast range of colours that could be produced by appropriate combinations of a small number of primaries led to speculations regarding the physiological basis of colour vision. Young also speculated that colour vision could be mediated by retinal mechanisms that responded selectively to each of the three primaries. In 1802 he advanced his theory of colour vision: "Now, as it is almost impossible to conceive each sensitive point of the retina to contain an infinite number of particles, each capable of vibrating in perfect unison with every possible undulation, it becomes necessary to suppose the number limited, for instance, to the three principal colours, red, yellow, and blue". Later he changed the principal or primary colours to red, green, and violet.

Newton had shifted the analysis of colour towards the physical dimension, while not excluding the subjectivity of colour perception. He also used the analysis of white light into its spectral components to synthesise new colours. For Greek scholars colour seemed to be a property of objects: it could be extracted from some plants and ores, in order to produce pigments that could be mixed with one another. Rough and ready rules for such mixtures were described by Plato, who recognised that the same pigments mixed in different proportions yielded different appearances. Some colours, however, could not be produced by mixing, and Aristotle noted that these corresponded to those seen in the rainbow. Experimental investigations of colour mixing started very early. Ptolemy devised a wheel on which different colours could be painted either as dots or in sectors;

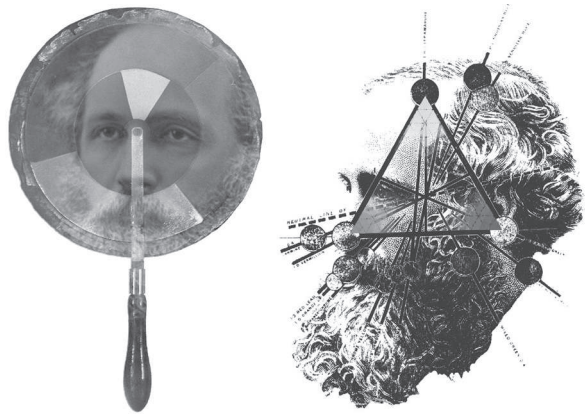


Figure 2.14 **Colour mixtures** James Clerk Maxwell (1831–1879) is shown on the left within a colour wheel of the type he used in his experiments. On the right, he is portrayed in combination with a diagram illustrating the chromatic relations of coloured papers; ultramarine, vermillion and emerald-green occupy the corners.

when the wheel was rotated the dots appeared as coloured circles or the sectors combined to produce a single colour: the persisting influence of one sector could combine with those of others. Colour wheels continue to be used to demonstrate colour mixture, and indeed were employed by James Clerk Maxwell (Figure 2.14) to provide a quantitative analysis of colour combination and apply it to the analysis of anomalies in colour perception.

Maxwell revived Young's three-colour theory, and also contended that there were three different colour receptors. Support for this derived from his studies of colour defective individuals, who could be classified according to the colour receptors that were absent or anomalous. Most colour defectives could match any colour with only two of the three primaries when using 'Maxwell's disc' for colour mixing. The resulting equations led to the proposal of a three-dimensional colour solid, with variables that would now be called hue, saturation and intensity. In 1861 Maxwell demonstrated the first colour photograph to a meeting at the Royal Institution in London. Separate photographs of a tartan ribbon were taken through red, green, and blue filters and projected through similar filters. The colour triangle had its origins in Young's work, but it was subsequently modified initially by Maxwell and then by Hermann Helmholtz (Figure 2.15).

Visual artists have been experimenting for much longer than visual scientists, and so it is natural that the practical skills that the painters had learned by trial and error should have had an influence on ideas about the nature of colour vision. The debate about primary colours was frequently based upon the practice of painting. Newton did distinguish between the mixture of pigments and that of light, noting that pigments reflected the incident light selectively. This could have proved useful to artists and scientists alike, but it was not pursued, perhaps because subsequent artists did not adopt the Newtonian primaries. The stimulus to differentiating light from pigment mixtures was Young's specification of a different set of primaries

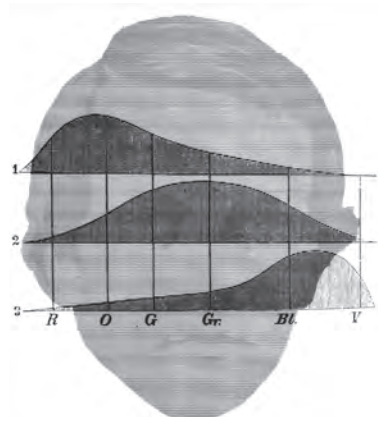


Figure 2.15 **Young Helmholtz** Hermann Ludwig Ferdinand von Helmholtz (1821–1894) extended Young's three-colour theory by presenting three speculative curves showing the respective wavelength sensitivities of nerves with these specific energies. A youthful Helmholtz, aged 26, is shown in combination with the curves depicting the analysis of white light into its three components.

for light (red, green, and blue) to those for pigments (red, yellow, and blue). The resolution was to await Helmholtz's clarification of the rules governing additive and subtractive colour mixing. When the three primaries for light are mixed white is produced, whereas the three pigment primaries combine to produce black; it is as if lights were added to one another, unlike pigments which subtract wavelengths. A pigment is so coloured because of the components of white light it absorbs and those it reflects; red, yellow, and blue pigments when mixed would absorb all the wavelengths in white light and so the mixture would appear black.

Representation

Colour is not the only dimension that has mixed the art and science of vision. Both are also intimately involved with the representation of space, as will be discussed in greater detail in Chapter 8. Pictures are such an integral part of our lives that we tend to overlook their remarkable peculiarities. We readily accept their relation to the objects that they represent even though they have few of the properties of those objects. Indeed, we rarely doubt the equivalence between the pictorial image and the object represented, and when this equation is questioned pictorially we are puzzled. The artist René Magritte painted a picture in the late 1920s which epitomised the problem of pictorial representation: it was called *The perfidy of images* and a variant of it is shown in Figure 2.16. The caption beneath the pictured pipe reads "This is not a pipe"! If it is not a pipe, what is it? The simple answer is that it is a picture of a pipe; the pictured pipe cannot be held or smoked, and it is not even supported by anything, as the object would need to be. The title chosen by Magritte for the work indicates that he was acutely aware

Figure 2.16 **This is not a pipe (after Magritte)** Magritte's picture was painted in oil on canvas whereas this picture was produced by photographing a curved briar pipe and copying Magritte's script beneath it.

of the problems associated with equating the pictures of objects with the objects themselves. Magritte was able to make this point forcefully because we can recognise the pictured pipe as representing a curved briar pipe. The orientation he chose for its representation was not arbitrary; had he painted it from other orientations its recognition would not have been as rapid.

Magritte was working within the style of representing objects in space (linear perspective) that has been adopted since the fifteenth century. Almost all Western artistic styles prior to the twentieth century were concerned with representing objects. As such they required a set of rules for generating the pictorial image – a set of relations that linked the object with the picture – which could be followed by the artist and understood by the spectator. Indeed, this was the very function of artistic representation. The fundamental problem that all pictorial artists must address is the difference between objects and pictures of them, namely depth. Over the last 40,000 years different solutions to this problem have been adopted, but all contain some common features that can be linked to optics. Science and art meet in perspective. Perhaps it should more accurately be said that the optics of antiquity met the art of the Renaissance in the context of linear perspective. The technique of perspective painting was rapidly adopted by artists from the fifteenth century onwards, and many textbooks described its rules (Figure 2.17). Linear perspective is now commonplace because photographic images follow the same optical principles on which perspective is based. However, there have been quite different styles of representing objects in space, and the objects represented can be readily recognised by us. The link between pictorial art and representation has been examined from the viewpoint of psychology by many writers.

Styles of representing space

The art historian Ernst Gombrich (1909–2001) referred to the different modes of representing space as "the riddle of style". He posed the question: "Why is it that different ages and different nations have represented the visible world in such different ways?" This was so despite the acceptance that there can have been little

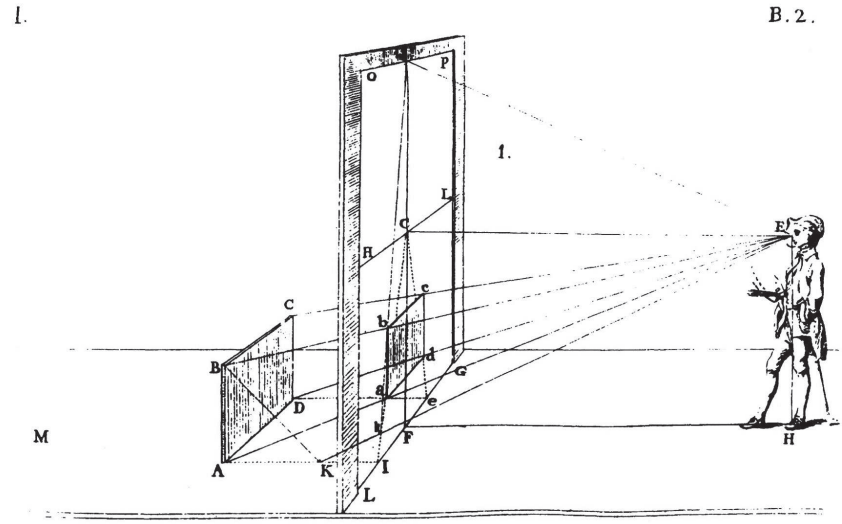


Figure 2.17 An illustration from Brook Taylor's classic eighteenth century treatise on perspective. "1. The Point of Sight, is that Point where the Spectator's Eye is placed to look at the Picture. Thus E is the Point of Sight. 2. If from the Point of Sight E, a Line EC is drawn from the Eye perpendicular to the Picture, the Point C, where the Line cuts the Picture, is called the Center of the Picture. 3. The Distance of the Picture, is the Length of the Line EC, which is drawn from the Eye perpendicular to the Picture. 4. If from the Point of Sight E, a Line EC be drawn perpendicular to any vanishing Line HL, or JF, then the Point C, where the Line cuts the vanishing Line, is called the Center of that vanishing Line. 5. The Distance of a vanishing Line, is the Length of the Line EC, which is drawn from the Eye perpendicular to the said Line: and if PO was a vanishing Line, then EJ will be the Distance of that Line. 6. The Distance of the vanishing Point, is the Length of a Line drawn from the Eye to that Point: Thus, EC is the Distance of the vanishing Point C, and EJ is the Distance of the vanishing Point J. 7. By Original Object, is meant the real Object whose representation is sought: and by Original Plane, is meant that Plane upon which the real Object is situated: Thus, the Ground HM is the Original Plane of ABCD."

genetic change in humankind over the last few millennia: the cave dwellers would have seen their world in much the same way as we do ours, although the contents of those worlds differed greatly. Objects would have been seen as having specific sizes and locations; they could be approached and grasped or avoided; some, like other animals, moved and their motion through space could be predicted. Survival would not be possible if we did not see the objects in the environment as constant despite changes in our positions with respect to them. One of the aspects that has not remained constant over time is how the objects are represented pictorially. Radical changes of style took place in the ways objects were depicted over the last 40,000 years of recorded art.

Gombrich has referred to the revolution in painting that took place in the early period of Greek art, during the sixth and fifth centuries BC when the desire to imitate nature was paramount. The artists strived to make the pictorial representations as lifelike as possible, so that, ideally, they would be confused with the real objects.

Figure 2.18 **Pioneers of perspective**. From left to right, Filippo Brunelleschi (1377–1446), Leon Battista Alberti (1404–1472) and Leonardo da Vinci (1452–1519).

In the first century AD Pliny related a story about a competition between two Greek artists in which one painted grapes so skilfully that birds tried to eat them, and the other painted a curtain that the first artist tried to draw open! Thus, artists attempted to make the pictorial image match the object so that it would deceive the eye. In this endeavour Greek painters would have had access to the developing science of optics, and basic aspects of projection, like foreshortening, could have been incorporated in their pictures. Most of the surviving examples are on pottery, although some mosaic compositions still exist.

The ancient texts on optics, when they were reintroduced to the West from the thirteenth century, were called *Perspectiva* and they treated direct vision, what Hero of Alexandria referred to as optics. The term perspective was derived from such texts, which often described aspects of pictorial representation and theatrical scene painting. However, perspective (or more correctly linear perspective) had a particular significance in the Renaissance because it reflected a return to the Greek pictorial ideal of attempting to imitate nature. The rules of perspective were devised in the intellectual cauldron of early fifteenth-century Florence. Linear perspective was demonstrated by architect and painter Filippo Brunelleschi and formalised by a contemporary mathematician, Leon Battista Alberti (Figure 2.18). Basically it was the application of Euclid's visual cone to a glass plane intersecting it, and this device is now known as Alberti's window. Thus, the principles of reducing a three-dimensional scene to a two-dimensional picture were formulated before the image-forming properties of the eye had been described. Nonetheless, the differences between looking at a picture and looking at a scene were clearly appreciated by Leonardo da Vinci (Figure 2.18). At the end of the fifteenth century he wrote "A Painting, though conducted with the greatest Art and finished to the last Perfection, both with regard to its Contours, its Lights, its Shadows and its Colours, can never show a *Relievo* equal to that of the Natural Objects". That is, the allusion to relief or depth in a painting of a scene, no matter how well it is painted, will be different to the perceived depth separating the actual objects in the scene.

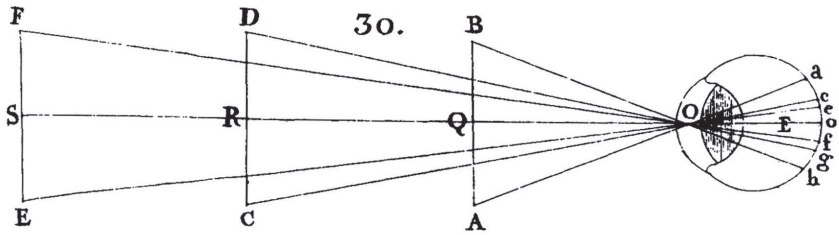


Figure 2.19 Objects of the same size, but at different distances from the eye, projecting differently sized images onto the retina, as was clearly described in Brook Taylor's treatise. "The farther distant the Eye is from an Object, so much less will the Picture of the Object be upon the Retina: for let E be the Eye viewing the several Objects AB, CD, EF at the Distance OQ, OR, OS. Having drawn the several Rays Aa, Bb, Cc, Dd, Ee, Ff, through the Pupil O, it will be manifest, that the Picture of the nearest Object AB, will be painted at the Bottom of the Eye in the Space ab, the Object CD in the Space cd, and the farthest Object, EF, in the Space ef."

Image formation

The influence of art was to prove significant because it framed a recurrent concern for theorists of vision: how can we perceive the world as three-dimensional when the image cast in the eye is two-dimensional? Isn't this precisely the problem that confronts the painter? Artists in the new perspective style seemed to answer this in practice if not in theory. An allusion to three-dimensionality could be induced if the rules of linear perspective were followed; this required a single, fixed viewing point and the depiction of objects in the scene in accordance with the angle they subtended at that viewing point. An object, say a person, will subtend decreasing angles at the eye as they walk away from us, even though they remain the same physical size (Figure 2.19). This change in projected size does not correspond to our perception; people appear to remain the same size as they walk away from us. Artists, in common with the rest of us, have a problem in depicting this state of affairs accurately because perception does not correspond to the dimensions projected onto the retina; hence they must use all manner of artificial devices, like matching the angles with an outstretched thumb, in order to discount their perception and record visual angles.

The idea of a particular view of a scene and a surface onto which the principles of optics could be projected is an integral part of linear perspective. Specifying a station point, picture plane, ground plane, and vanishing point permitted depiction of a single image of a scene. Alberti described the rules for capturing the image, and a number of techniques for representing it. Some of these were mathematical, others involved a grid through which the scene is observed and a similar grid onto which the lines can be drawn. A famous woodcut by Albrecht Dürer (1471–1528) shows an artist painting a model using such a system. A gnomon or sighting vane was used in order to define a station point, and the figure was viewed with the eye as close to it as possible. Perhaps the simplest means of drawing in perspective

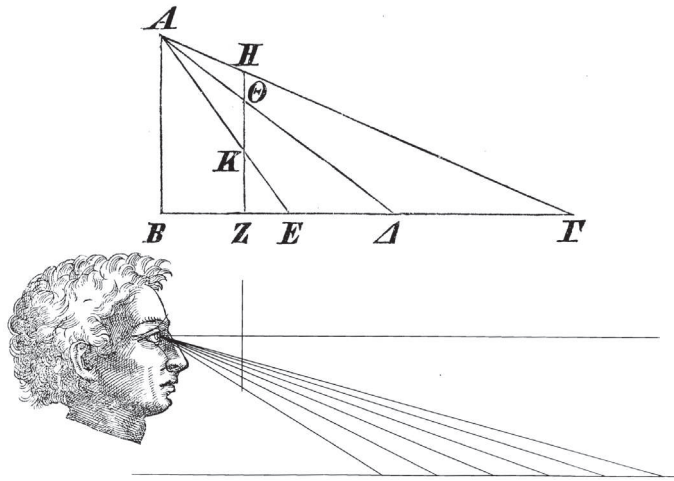


Figure 2.20 Upper, an illustration from Euclid's *Optics*. Lower, a representation of Alberti with his window; the vertical line is equivalent to the window or picture plane.

was 'Alberti's window'. If a single eye is maintained at a fixed position with respect to a window then the objects in the scene beyond the window can be represented in accurate perspective simply by tracing their outlines on the surface of the window. An illustration of Alberti's window is shown in Figure 2.20, together with a similar figure from Euclid's *Optics*. Another device, which combined the science and art of optics, was marshalled by some artists to form images in accurate central perspective – the camera obscura.

Camera obscura

The general principles of image formation have a long history. Shadow casting was known to scientists in China as early as the fifth century BC, and the practical optics of the camera obscura, or dark chamber, were described by Chinese scientists in the ninth century. Little in the way of a theory of image formation is considered to have derived from these experimental enquiries. Ibn al-Haytham described an inverted image in a dark chamber, and pinhole experiments were conducted by late medieval students of optics. However, the equation of such optical image-forming devices with the eye appeared much later.

Leonardo da Vinci conducted experiments with a camera obscura, and drew an analogy between its operation and that of the eye. In this regard, as in many others, his ideas were neither published nor widely known until long after his death. A later sixteenth-century artist, Daniele Barbaro (Figure 2.21), appreciated the assistance that the camera obscura could offer to the painter, particularly when a convex lens was placed in an enlarged aperture. A powerful convex lens was recommended, and the sharpest image could be located by moving a sheet of paper towards or away from the aperture. He also realised that the sharpness of

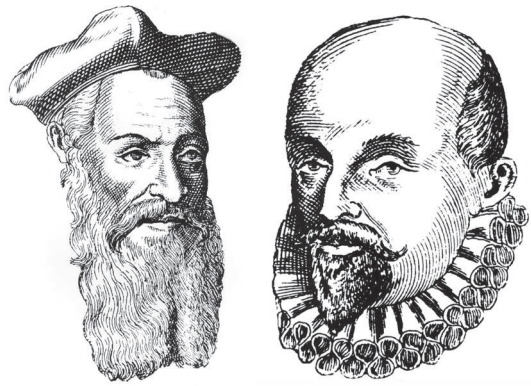


Figure 2.21 Both Daniele Barbaro (1513–1570), left, and Giovanni Battista della Porta (1535–1615), right, saw the benefits that the camera obscura could provide for artists, particularly when a convex lens was placed in an enlarged aperture. Porta applied the analogy of the camera obscura to the eye.

the image could be further enhanced by reducing the size of the aperture through which the light entered the dark room.

The camera obscura was used by both artists and scientists to generate images. The artists realised that the tedium of perspective projection could be circumvented by its application, and scientists considered that it offered a solution to the problem of vision. For example, Giovanni Battista della Porta (Figure 2.21) argued that the projection of images in a camera obscura provided conclusive evidence against an extramission theory of vision, and that it explained "how vision is made"! Porta still held to the belief that vision took place in the lens. The link between anatomy and optics was still to be forged.

Development of perception

All aspects of perception and behaviour undergo developmental changes. That is, the range of function changes with age. Perceptual development can be distinguished from perceptual learning since the latter is not necessarily related to age. Perceptual learning is a vital part of adaptation to our environment and is so pervasive that we can easily overlook its importance. We are constantly learning to discriminate between stimuli that initially appear similar. Our ability to recognise so many individual faces is an example of this. All faces have many features in common, but we are able to determine the identities of particular faces because of the slight differences in the configuration of the parts and the ways in which they move to express emotions. Children engage in perceptual learning when they grapple with the oddities of the patterns we use in written language.

Before describing some of the developmental changes that occur throughout the lifespan, something will be said about measuring perception in infants and animals. Experiments on perception communicate the experimenter's requirements

to an observer and measure some subsequent response. For psychophysical experiments with humans this typically involves communicating the requirements with language and indicating how responses will be made. But what about cases where this cannot be done? These procedures could not be applied to determining the perceptual capabilities of our metaphorical guide dog or to other animals. In addition, we may wish to investigate perception in infants or perhaps in people whose linguistic abilities are impaired. For these, the requirements of the experimenter must be communicated by some other means, and the response cannot be verbal. Clearly, it must be some action which is within the behavioural repertoire of the observer. Thus, for example, it may be possible to use the methods of conditioned learning to study discrimination between stimuli. If one response, like turning the head or raising a paw, can be conditioned to a red stimulus, and another to a green one, then we may be justified in concluding something about the ability to perceive colour; provided of course that the discrimination is not based on some other characteristic like the brightness of the stimulus. If an infant spends more time looking at a picture of a face than at a random collection of lines, then this preference may demonstrate an ability to recognise faces as a special class of object. Clearly it would also be necessary to establish that these measurements did not simply reflect a preference for symmetrical patterns, or even for looking left rather than right. Such behavioural measures are not intrinsically different from verbal ones, and similar sorts of inference may be made from them. However they are less subject to biases, and the inferences are therefore likely to be more secure.

It is possible to be seriously misled by measurements of perception, due to their simplified and inferential character. On the one hand, we might consider that more detailed discriminations are available to an observer than actually exist, and this can apply particularly to animal perception. Conwy Lloyd Morgan (Figure 2.22) issued a canon of parsimony to guide interpretations of animal perception and behaviour: "In no case may we interpret an action as the outcome of the exercise of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one which stands lower in the psychological scale." Morgan was reacting to the use of anecdotes which imputed high levels of reasoning to pets. His counter to anecdotalism was a more keenly observed anecdote regarding the ability of his fox-terrier to lift a latch on a garden gate. "Now the question in any such case is: How did he learn the trick? In this particular case the question can be answered, because he was carefully watched. In this case the lifting of the latch was unquestionably hit on by accident, and the trick was only rendered habitual by repeated association in the same situation of the chance act and happy escape. Once firmly established, however, the behaviour remained constant throughout the remainder of the dog's life." He referred to the initial activity of his dog as trial-and-error learning, and argued that apparently intelligent behaviour can only be understood if its development has been investigated. The same conclusions apply to perception.

On the other hand, we may suppose from observable behaviour that perception is limited or even non-existent, when in fact this is not the case. Another example from animal behaviour can be used to make this point; it concerns Clever Hans –

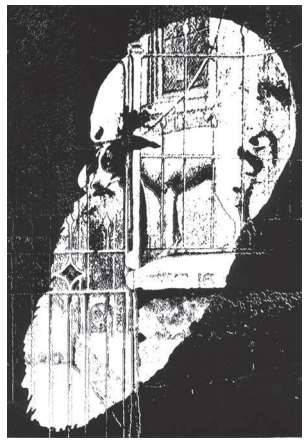


Figure 2.22 **Lloyd Morgan's canon** Conwy Lloyd Morgan's (1852–1936) portrait, with his long, flowing beard, is presented together with a drawing of his fox-terrier lifting the latch of the garden gate. Observation of the dog learning to perform this action was the basis for Lloyd Morgan's canon of parsimony.

a horse that was reputed to be able to count and perform arithmetic calculations. When the horse's owner asked questions regarding, say, addition Hans would tap the answer with his hoof. These remarkable feats from a century ago were shown to be due to the acute perception of the horse: when the correct answer was approaching, changes in the owner's posture and expression were detected. The owner was unaware that cues were being given in this way, but when the answers were required without sight of the owner the horse lost its remarkable powers. The commission assembled to assess the claims for Clever Hans essentially applied Lloyd Morgan's canon – they sought the most parsimonious interpretation of the behaviour that was observed. They also uncovered aspects of horse visual perception that had not been appreciated previously.

The traditional view of perception in early infancy was that human babies could see very little, and recognise essentially nothing. Only with the development of better methods of measurement, more closely related to the behavioural repertoire of infants, has it been shown that their perceptual abilities are in fact quite considerable. This finding has been important for the early diagnosis of sensory defects, and has given new significance to the potential effects of neonatal perceptual experience.

Perceptual development is not restricted to the infants of species. They apply both to the structures of the visual system and the functions that they perform. There are changes in vision over the whole life span, and particularly systematic ones in the later decades of life. One obvious and well-known structural change is that associated with the ageing eye; the lens loses much of its elasticity, so that it cannot focus or accommodate over the same range. Near objects cannot be resolved nor can those at a distance. Fortunately, these developmental changes in the optical functioning of the eyes can be compensated for by wearing specially prescribed

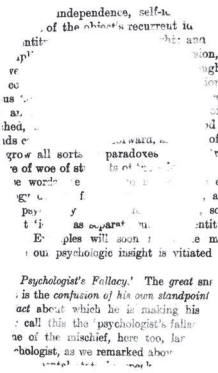
eye glasses: these are said to have been amongst the most significant inventions in human history. Other changes are more subtle, and often go unnoticed.

Infant vision

One question concerning infant vision coursed through philosophy long before it was examined in psychology – what can newborn infants see? The question was of interest to philosophers because it addressed an old issue: to what extent is our knowledge about the world a consequence of our experience? Rationalists considered that it is dependent on thought processes rather than the senses, whereas the empiricists argued the reverse. Accordingly, the perceptual experience of newborn infants offered a way of addressing these conflicting philosophies but there were no suitable means of testing it until recently. An alternative approach, which was open to study, was to examine vision in individuals who had been blind from birth and then had sight restored (usually by removing cataracts). William Molyneux (1656–1698), an authority on optics and astronomy, had a particular interest in this problem as his wife was blind. He famously stated the question in a theoretical way, and also suggested an answer to it. It is now frequently called Molyneux's question, and it was posed to the empiricist philosopher, John Locke:

> Suppose a Man born blind, and now adult, and taught by his touch to distinguish between a Cube and a Sphere of the same metal, and nighly the same bigness, so as to tell, when he felt one and other, which is the Cube, which the Sphere. Suppose then the Cube and Sphere placed on a Table, and the Blind Man to be made to see. Quære, Whether by his sight, before he touch'd them, he could now distinguish, and tell, which is the Globe, which the Cube. *To which the acute and judicious Proposer answers:* Not. For though he has obtain'd the experience of, how a Globe, how a Cube, affects his touch; yet he has not yet attained the Experience, that what affects his touch so or so, must affect his sight so or so; Or that a protuberant angle in the Cube, that pressed his hand unequally, shall appear to his eye as it does in the Cube.

The answers of Molyneux and Locke were negative on theoretical grounds, but some years later postsurgical evidence was brought to bear on the argument. The most celebrated early operation was performed by William Cheselden (1688–1752) in London. He restored the vision of a boy who had been blind from childhood due to cataracts and carried out a number of informal tests on the vision of the operated boy in order to determine what could be discriminated. The distances, sizes, and shapes of objects could not be differentiated. Moreover, pictures of objects provided particular problems for perception, and it took about two months before they were recognised as representations of other objects. There was never any question amongst the philosophers about whether the person with sight restored would be able to see postoperatively, but only whether they could name objects by sight alone. Physicians, on the other hand, were faced with the



ndependence, self-h
of the object's recurrent id
ntit- hic and
ion, s
ugh a l
ions. T
ed, an
nstru
d intro
of whic
radictio
paradoxes
e of woe of st
e t
, a
so
ntit
e m
oun psychologic insight is vitiated

Psychologist's Fallacy.' The *great* sin
is the *confusion of his own standpoint*
act about which he is making his
call this the 'psychologist's falla
ne of the mischief, here too, lar
chologist, as we remarked abov

Figure 2.23 **The psychologist's fallacy** William James (1842–1910) is shown in text from his *Principles of psychology* describing the psychologist's fallacy: "the assumption that the mental state studied must be conscious of itself as the psychologist is conscious of it".

practicalities of vision in those with sight restored. In the early nineteenth century, attention was drawn to the uniqueness of Cheselden's case, and the difficulties they had in making similar general statements from their own experiences. Molyneux's question has stimulated investigations many times since it was first enunciated. It has resurfaced recently in the remarkable case of a man who received stem cell transplants to restore his sight which was lost in a childhood accident 43 years previously. His recovery was charted with the armoury of modern psychophysics and imaging methods. As with most cases of restored vision, the blessings of the new sense were mixed: the low vision that was available was often competing with well learned non-visual strategies. Despite the insights derived from contemporary investigations into the question posed by Molyneux, the answer he proposed remains questionable.

With the emergence of experimental psychology in the late nineteenth century more attention was directed to the perception of infants, although the interpretations were not based on experiments. William James (Figure 2.23) remarked, from his armchair, that "The baby, assailed by eyes, ears, nose, skin, and entrails at once feels it all as one great blooming, buzzing confusion". The conclusion concerning infantile confusion was based on the observation of their seemingly random and chaotic movements. More detailed scrutiny indicated that some aspects of behaviour were systematic and could be used to determine what interested infants. Many novel methods have been devised to study infant vision and they have yielded knowledge about the developmental course of vision and visual discrimination. For example, it was noted that infants spent different amounts of time fixating on visual patterns and so infant perceptual discrimination was inferred from the patterns of preferential fixation. Recordings of eye movements in infants only a few days old showed that they were concentrated on contours or corners of simple patterns. When a stimulus is presented many times the response to it typically declines or habituates.

Habituation to repeated presentations of patterns provided another source of inference regarding discrimination, particularly when novel patterns were presented; if the infants dishabituated then it was assumed that the novel pattern was discriminated from the habituated pattern. Operant conditioning techniques were applied to demonstrate the emergence of perceptual constancies. These methods were refined and the course of perceptual development began to be charted. Now the visual development of infants can be assessed with physiological methods like ERPs and MRI. Moreover, detailed neurophysiological studies of the developing visual systems of closely related species (like macaque monkeys which have similar visual anatomy to humans) can be related to the changes that take place in human infants. Such investigations have pointed to the importance of early visual experience in determining the later functioning of vision.

Many of the features of visual development in infancy depend upon changes taking place in the visual nervous system. For example, the retina itself is not fully formed at birth and of the two types of receptors (rods and cones) the population of cones (which are concentrated in the region of fixation and which mediate colour vision) is still increasing. The cones become more numerous, thinner and more densely packed. Accordingly, colour vision in newborn infants is rather poor, but it does improve rapidly in the first weeks and months of life. In order to extract the spatial detail from an object it needs to be focused on the retina and followed if it is moving. The state of focus (accommodation) will need to change for objects at different distances. At birth focus is fixed at around 20 cm but from two weeks to three months accommodation improves to be similar to that in adults. Indeed, the receptors in the retina are not fully developed at birth nor are the nerve cells in the visual cortex, and so their development is likely to have a profound effect on what can be seen. It is not surprising, then, that the visual acuity of the newborn is more than ten times poorer than that of adults, but it improves rapidly in the first months of life until it reaches almost adult level at age six months. Similarly, infants in the first few months of life are not able to detect low contrast patterns (where the differences between the lightest and darkest parts are small) that are readily detectable by adults. Another distinguishing feature relates to the ways the eyes move. Adults can follow slowly moving objects (called pursuit eye movements) as well as flick quickly between different parts of objects observed. The flicks or jerks are called saccades; they occur about three times a second and are punctuated by periods of fixation. At birth, infants perform saccades the dynamic characteristics of which become similar to those of adults at around three months. Smooth pursuit movements are lacking in the newborn and appear in the third month of life. One aspect of eye movements that is present at birth is the reflex following of a moving object; it is called optokinetic nystagmus because the eyes move slowly in one direction and then make a ballistic return and repeat this sequence. The two eyes do not move as a unit at birth and are integrated gradually from the second month. Binocular alignment is a prerequisite of stereoscopic vision and the latter develops in the following months, enabling the infant to discriminate between objects at different distances.

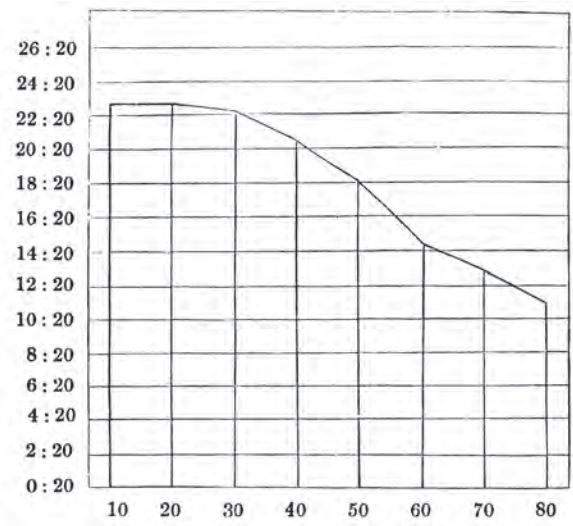


Figure 2.24 Normal visual acuity is given as 20:20 and this chart from the mid-nineteenth century indicates the declining acuity with age.

Aged vision

Ageing is not a novel condition and so visual changes associated with it have a long observational history. Perhaps the most readily evident change is the restriction in the range of accommodation, or presbyopia. Such difficulties of vision in old age were described long before any means of compensating for them had been adopted, and these in turn were applied long before the reasons for their efficacy were understood. The glimmerings of knowledge about these topics appeared towards the end of the sixteenth century when a convex lens was placed in the aperture of a camera obscura and such a camera was likened to the eye. With the clearer understanding of the optics of presbyopia, corrections for it became routine, although doubts remained concerning its causes. Thus, convex lenses corrected the defect, resulting in distinct vision for nearer objects. Myopes, who required concave lenses in their youth, were able to see distant objects more distinctly in old age; this demonstrated an appreciation of the developmental interaction of myopia and presbyopia. The developmental course of presbyopia was intensively studied in the early nineteenth century and it was represented graphically (Figure 2.24). Modern techniques can incorporate variations in the sign and strength of corrections in external lenses, in contact lenses and in lenses implanted in the eye itself.

Reductions in the range of accommodation are not the only changes which occur in old age. The sensitivity to light decreases and the range of colours that can be discriminated declines. The lens thickens and yellows with age, both reducing the amount of light striking the retina and changing its spectral composition:

more short-wavelength (bluish) light is absorbed by the lens. There is also the greater likelihood of certain ocular diseases, like cataract (when the crystalline lens becomes cloudy), glaucoma (when the pressure within the eye increases) and macular degeneration (where the cells in the sensitive region around the fovea die), which can have drastic visual consequences. In addition, with advancing age it becomes more likely that floaters will be seen – small, opaque and moving spots. These have been given some graphic description in the past, like flying gnats, because they appear to an observer as small dark blobs which seem to move about as the eyes do. They are seen most readily when a clear background is observed, like a white wall or the blue sky. The retina and its associated structures are in constant growth and decay and the floaters represent the visibility of accumulations of discarded cells.

Illusions

Despite the long history of recording perceptions, attention continues to be directed to the minor variations that we call errors of perception or illusions. The term perceptual error is a very strange one. There can only be perceptions. Errors are associated with some deviation from a reference or standard. Thus, an error of measurement refers to some deviation from a reading that can be shown (either by another instrument or by other observers) to be discrepant. By the same token, the term misperception is a misnomer. However, there is some utility in the term if the dimensions of time and space are incorporated. A discrepancy in the perception of the same object can occur over time. In like manner, the colour of a surface can be modified by surrounding it by one of a different colour. In these cases, perceived object properties are variable. An illusion requires a yardstick or reference relative to which it can be assessed. In fact, underlying virtually all illusions are the mismatches of observation described above. However, in terms of interpretations the interest is applied to the source of the mismatch. A classical example is the apparent bending of a stick when immersed in water. There is a compelling contrast between the appearance of the stick in air and partially immersed, but does the stick change its characteristics? The investigation of aspects of optics derived from precisely such observations.

Errors in perception were remarked upon before the basic perceptual processes were either described or appreciated. This was so because it was possible to compare observations of the same objects over time and to note any discrepancies between them. The modern definition of illusions applies to differences between the perception of figures and their physical characteristics. Consensus concerning an external reality did not exist in antiquity, and so attention was directed to those instances in which changes in perception occurred. That is, when the same object appeared to have different properties under different conditions. According to this observational definition of illusions, all that is required is an assumption of object permanence; thereafter, any changes in the appearance of the same object will be classified as illusions.

Figure 2.25 A nineteenth-century engraving of the lower Fall of Foyers; the River Foyers runs into Loch Ness in northern Scotland. Addams would have observed the waterfall from the observation point shown.

Visual illusions

An abiding example of this is the variations in apparent sizes of celestial bodies. The moon illusion – its larger appearance near the horizon than high in the sky – is a size illusion, but it has also been interpreted as a distance illusion. The moon illusion presented an enigma in the past and it is one that still persists. Modern attempts at explaining it remain problematical. It provides a quintessential example of illusion because the observations have been consistent but the interpretations have shown a progressive change: It was analysed initially as a problem of physics, then physiology, and finally psychology. The moon illusion was known to Aristotle, as was another striking visual illusion that occurs in the natural environment, which has been called the waterfall illusion, but it is now referred to as the motion aftereffect (it is also referred to as the MAE).

In the third century BC, Aristotle described the apparent motion seen in stationary stones at the riverside after he had observed pebbles beneath the flowing water. Little more was written about the illusion until the nineteenth century. Jan Evangelista Purkinje (1787–1869) observed a cavalry parade for more than an hour, and when he looked at the houses directly opposite they appeared to move in the reverse direction to the parade. A few years later Robert Addams (c. 1800–1875) observed the Falls of Foyers in northern Scotland (Figure 2.25) and described the waterfall illusion thus: "Having steadfastly looked for a few seconds at a particular part of the cascade, admiring the confluence and descussation of the currents forming the liquid drapery of waters, and then suddenly directed my eyes

Figure 2.26 **Plateau spiral** Joseph Antoine Ferdinand Plateau (1801–1883) was a pioneer in the study of visual persistence. In 1832 he invented a means of presenting stimuli briefly and in sequence (the phenakistoscope) to yield apparent motion. He also examined the motion aftereffect using a spiral pattern like the one in which his portrait is embedded.

to the left, to observe the vertical face of the sombre age-worn rocks immediately contiguous to the water-fall, I saw the rocky face as if in motion upwards, and with an apparent velocity equal to that of the descending water."

This was only one of the rediscoveries of the motion aftereffect: Joseph Plateau (Figure 2.26) stumbled across it in the course of conducting experiments on apparent motion. He had been studying the effects of rotating patterns on perception and noticed the motion visible in stationary objects following prolonged exposure. He introduced the stimulus employed most widely throughout the second half of the nineteenth century – a black disc with a white Archimedes spiral on it – which became called the Plateau spiral. The earliest interpretations of the waterfall illusion were in terms of eye movements despite the logical difficulty of accounting for the phenomenon in this way: any after-movement of the eyes would affect all contours in the visual field, rather than being restricted to regions exposed to prior motion. The illusion continues to be studied because it can now be related to motion processing in the brain (see Chapter 7).

Assuming that the physical world is stable, then an illusion occurs when the same object appears to have different properties (of colour, position, size, shape, motion, etc.) under different circumstances. The changing perception of objects was the source of Aristotle's interest in phenomena like afterimages, aftereffects, colour contrasts, and diplopia (double vision). It was precisely such variation in perception that led Plato to a distrust of the senses. Many commentators have argued that illusions are a modern preoccupation in the study of perception. This statement, when restricted to geometrical optical illusions, is certainly correct

because they only received this name in the mid-nineteenth century. Moreover, the phenomena so called are often associated with a particular theoretical outlook, based on laboratory studies of simple stimuli whose characteristics can be readily controlled. Illusions in the natural environment may be observed and described, but do not lend themselves to systematic study. Students of vision have been more intrigued by illusions of the senses than by the veridical perception of object properties. Since most of perception is veridical, it is the occasional departures from veridicality that have provided fascination. Illusions, or errors of perception, have been gauged by many means. The most venerable method has involved the comparison of percepts. When object properties appear to differ under different circumstances, then an illusion is said to have occurred. The only assumption that needs to be made is that the objects have not themselves changed between the two events. Physical benchmarks were introduced along with an increasing understanding of the physical world. Thus, perplexing percepts like the apparent bending of sticks when immersed in water could be given a more mundane interpretation when the laws of refraction were specified. However, the reference that has proved most attractive to many students of the senses is comparison between the illusory percept and the underlying physiology. This attraction has only been increased by more sophisticated means of studying the physiology, but the hypothesis that illusions reveal the normal working of perception is still essentially untestable.

For many centuries, magicians have been the masters of illusion in the everyday world. Probably most people's first contact with illusions is through magic, and it certainly reaches a larger section of the population than the illusions studied in science. We now accept that magic involves tricks that comply with natural laws, but this was not always so. Magicians of the past appreciated the laws of light and could manipulate attention with far greater subtlety than was the case for students of the senses. They were able to beguile and bamboozle those who did not share their knowledge, which they kept as secret as possible. Magicians use their practical knowledge of illusions and selective attention with consummate skill, without delving too deeply into their underlying basis. The second half of the nineteenth century was the era of the scientific study of illusions. Despite the fact that illusions have fascinated students of the senses for over 2,000 years, many of the familiar geometrical optical illusions that are graced by the names of their discoverers derive from this period. The optical illusions devised in the nineteenth century are not so magical, but they are robust and they have had an enormous impact on the development of psychology and the understanding of perception.

There are many examples of spatial illusions, and a long history of experimentation and theory designed to account for them. It is important to note that the idea of a visual illusion presupposes that the object or pattern concerned would be perceived differently under other conditions. For example, the apparent length of a line is altered by adding oblique lines, known as fins, to either end or the orientation of parallel lines appears distorted because of the angles at which they intersect others (Figure 2.27).

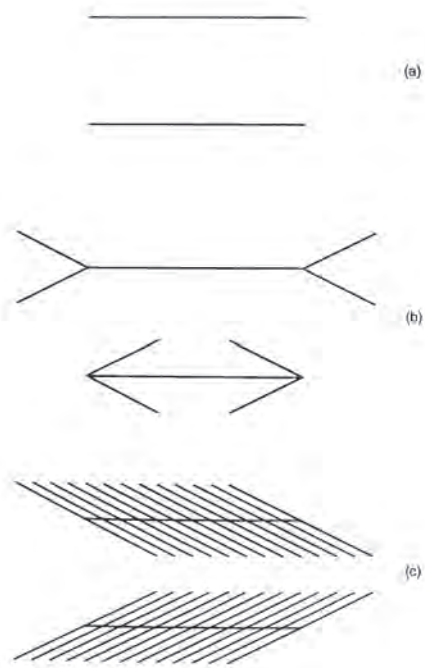


Figure 2.27 (a) Two equal and parallel lines. (b) The same lines with fins (Müller-Lyer illusion). (c) The same lines crossing diagonal lines (Zöllner illusion).

The Müller-Lyer and Zöllner illusions are instances of the distortion of the perceived geometry of simple plane figures – the former of size and the latter of orientation. That they are illusions can only be shown by measuring the lines with a ruler (which assumes correct perception of the ruler itself), or by comparing the perceived lengths of the lines without the inducing parts. Essentially, illusions are defined by comparisons between perceptions, although we naturally assume that at least one of these is physically correct. A more compelling way of demonstrating illusions is by presenting the component parts separately so that they can be combined. In the case of the Müller-Lyer figure, the horizontal lines can be shown in isolation and then together with the fins.

One view of illusions is that they can be used as tools to probe the mechanisms of visual perception, because perceptual errors can give us clues about the way in which normal perception takes place. For example, the Müller-Lyer illusion has been accounted for in terms of depth due to perspective. If we interpret two-dimensional drawings as representing an object in depth, then there may be systematic distortions of apparent size. This would be due to the process which ensures that real objects appear to remain their actually constant size despite being at different distances, and projecting different-sized images on the retina. A normal perceptual process would thus have been revealed by its inappropriate influence on a line drawing, and it could be studied by this means. One problem

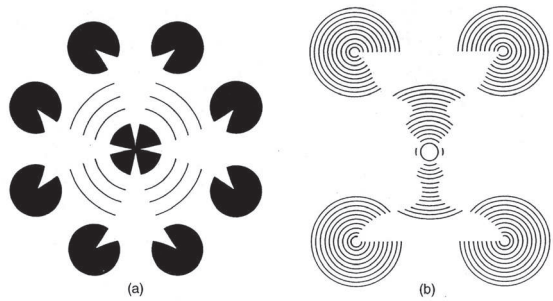


Figure 2.28 (a) Illusory Maltese cross. (b) Illusory vase/faces figure. The appearance of contours where no luminance boundaries exist has also been called illusory, subjective, cognitive and virtual contours.

is that many visual illusions appear to be the result of several different perceptual processes, and it is hard to separate their respective effects.

Visual spatial illusions are generally classified as distortions of orientation or size in simple line patterns. Visual neuroscience is generally concerned with neural and/or behavioural responses to contours. For this purpose, stimulus contours are usually defined as luminance discontinuities, and these can be enhanced by interactions at early stages of visual processing. However, one set of stimuli involves seeing discontinuities of brightness where no luminance differences exist. These illusory contours can be produced by gaps in shapes, as in Figure 2.28a. The white triangular shapes that radiate from the centre appear brighter than the white background. The shapes are produced by the white sectors of the dark discs and by the terminations of the concentric circles. The illusory contours even appear to curve between the discs rather than follow the shortest straight line. Illusory contours can also be induced by line terminations as in the case of the central vase shape of Figure 2.28b. Illusory contours behave much like physical ones – they interact with one another and can modify the appearance of subsequently presented patterns, producing spatial aftereffects.

Other pictures are far more paradoxical. For example, the joint in Figure 2.29a appears very strange – almost impossible. However, it is based upon a simple perspective principle: a rectangular rod can be represented by three parallel lines. The central quadrilateral can represent the end of either an upper or lower rectangular rod. Each would be perceptually plausible in isolation, but their connection creates a paradox. A junction between two rectangular rods can be depicted by six lines, with the depth suggested by the occluded edge. A cylinder, on the other hand, can be described by only two lines and an ellipse, so the lines showing several rectangular junctions would make more cylinders. The central 'impossible' triangle in Figure 2.29b has four rectangular rods extending from each side which are transformed into six cylinders at the extremities of the lines. These are called impossible figures because the solid objects to which they allude could not be constructed. Impossible worlds can be drawn by manipulating

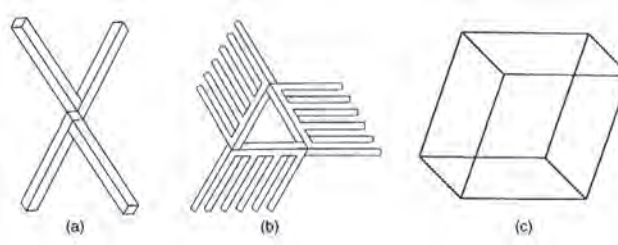


Figure 2.29 Paradoxical figures: (a) Impossible junction. (b) Double impossibility. (c) Reversing figure.

the rules of perspective, and artists like M. C. Escher (1898–1972) have been exceedingly skilful at producing perceptual paradoxes of this type.

These are not, of course, impossible objects (as they have been called at times) but perfectly possible pictures. Pictures provide us with allusions to objects, and tricks can be played with the transition from three to two dimensions. Indeed, this is grist to the artistic mill: pictures incorporate ambiguities that are rarely present in objects. One such ambiguity is the depth that is represented in simple line drawings like Figure 2.29c: the apparent orientation of the rhomboid changes so that the front face appears to be pointing either down and to the right or up and left. The picture is interpreted as representing a three-dimensional structure, but there is insufficient detail (for occluding contours, perspective convergence, or texture) to define which parts would be near and which far. However, we do not entertain both depth interpretations simultaneously: our perception flips from one possibility to the other, so that the apparent depth undergoes reversals. Perhaps the most familiar form of ambiguity relates to drawings which can be interpreted in more than one way, because the same contours delineate two different figures, like the vase/faces motif in Figure 2.28b.

It is, however, instructive to examine why there should have been this burst of illusory activity in the late nineteenth century. Put another way, why did outline drawings assume such a central role in the study of perception? It could have been due to the combination of two powerful strands of thinking about vision. The first stems from the seventeenth century, when the dioptrical properties of the eye were elucidated, thereby setting in train the idea that the problem of perception has as its starting point the static, two-dimensional retinal image. The retinal image was considered as static, and the problem was seen as restoring the missing dimension of distance from the ambiguous projection. The second strand relates to the experimental approaches to the study of perception introduced in the mid-nineteenth century. It was argued that the experimental rigour of the physical sciences should be brought to bear on the study of perception. Thus stimulus variables should be isolated and manipulated in quite unnatural ways in order to determine how perception is modified. It is difficult to manipulate solid objects, but it is exceedingly easy to create novel pictures. So pictures became the accepted stimuli for the study of vision. Once accepted, the psychologists then

rediscovered tricks that had been a part of the artist's armoury for centuries and they devised some novel ones, too.

Are pictures so central to understanding perception? What is the relationship between the perception of pictures and of the objects they portray? Will understanding pictures facilitate our interpretations of vision, or vice versa? And how do those peculiar pictures, geometrical optical illusions, relate to other forms of pictorial representation? It could well be argued that the study of perception will not be furthered by the examination of such oddities, and such arguments are presented in Chapter 8.

Purkinje's images

The distinction between illusions and hallucinations has often been difficult to draw. One way of separating them is to refer to the population experiencing the phenomena. Generally speaking, visual spatial illusions are seen by everyone, albeit with different magnitudes. However, if someone says that they can see an elephant in the room, this perception may not be shared by others. Thus, hallucinations can be thought of as idiosyncratic rather than universal. This simple distinction breaks down when we consider some of the oldest descriptions of states that were called hallucinations – they were referred to as subjective visual phenomena by Purkinje. Despite the fact that many of the subjective phenomena he examined had previously been described, his contribution was to classify the observations, to describe them with precision, and interpret them within the context of the then known physiology. Purkinje's representations of these effects are shown in Figure 2.30. He described stroboscopic patterns, pressure figures, effects of galvanic stimulation, blind spot, pattern disappearances and distortions, visibility of retinal blood vessels and blood flow, afterimages, aftereffects, single and double vision, eye movements, peripheral vision, real and apparent motion, accommodation, near and far vision, and voluntary squinting. We can take stroboscopic (flickering) patterns as an example of his approach: he produced flicker by waving his fingers in front of one eye, while looking at the bright sky, and reported seeing checkerboards, zigzags, spirals, and ray patterns. They are now referred to as stroboscopic patterns although they were described before the stroboscope was invented. When the eye is stimulated by an unpatterned, flickering light, patterns of bewildering complexity become visible. They are called stroboscopic patterns after the modern form of the instrument, which can deliver pulses of light at very high frequencies.

Purkinje also described a range of distortions that can be seen both during and after prolonged observation of geometrical patterns like parallel lines and concentric circles (Figure 2.30). Stationary concentric circles provide a different type of distortion to parallel lines: cloudy streaks appear to radiate from the centre, and they can themselves rotate. Purkinje associated these effects with variations in accommodation, although he did not say how they were related. The distortions were radial with concentric circles and circular with a pattern of radiating lines, that is, they were perpendicular to the lines in the pattern. Such effects can be

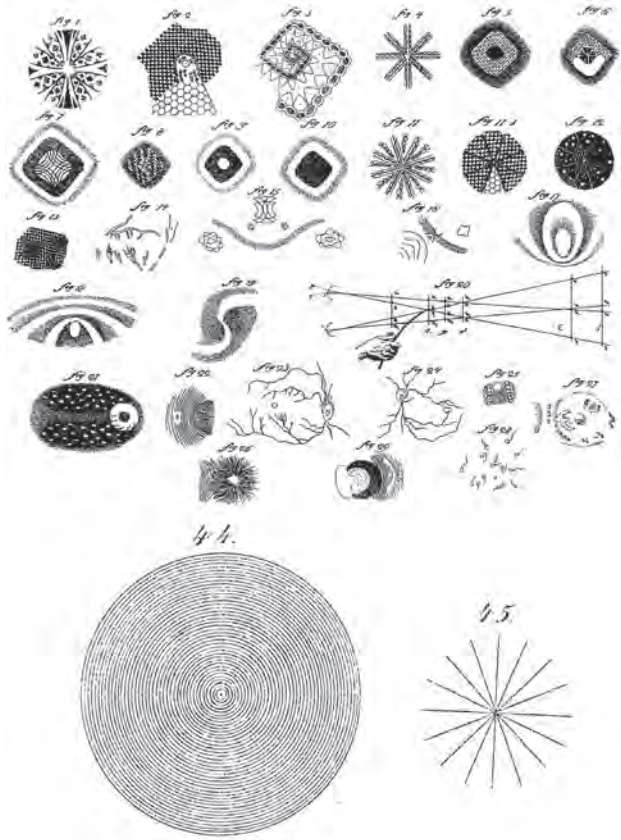


Figure 2.30 Upper, a plate from Purkinje's first book on subjective visual phenomena. Figures 1–4 were light and shadow figures, 5–14 were patterns visible when pressure was applied to the eye, 15–16 were experiences seen when the eye was stimulated electrically, 17–19 were cloudy streak patterns, 20–22 were concerned with determining the blind spot and the luminous circles visible around it, 23–24 were the shadows of retinal blood vessels cast by light at the corner of the eye; this pattern is known as the Purkinje tree and 25-29 represent the visibilities of entoptic phenomena that Purkinje called blood circulation in the eye, curvilinear radiations, pulsating figures, flying gnats and luminous rings, respectively. Lower, two figures from his second book showing patterns of concentric circles and radiating lines; these appear to become distorted with prolonged viewing and they also generate dramatic aftereffects when a blank surface is subsequently viewed.

attributed to transient forms of astigmatism that are a consequence of instabilities in the ciliary muscles.

Purkinje was not content to describe these phenomena but he wished to provide interpretations of them, and saw them as a key to understanding perception generally: he said that visual illusions reveal visual truths. The pattern distortions visible during or after observation of geometrical patterns were studied by him for the first time. A variety of such distortions are now known to occur; the effects

and aftereffects were not only described by Purkinje but he also attempted to account for them by manipulating the conditions under which they can be seen. For example, he found that the aftereffects were dependent upon the line spacing (spatial frequency) of the gratings. The perceptual transformation of parallel lines into wavy ones was accounted for by the overlap of real images with slightly displaced afterimages, like optical interference fringes. Thus afterimages assumed an interpretive as well as a phenomenological role. There has been a resurgence of interest in these phenomena as a consequence of their use by artists and their examination by visual scientists.

The shimmering, back-and-forth motion over a blank field is rather like a motion aftereffect but there is no physical motion that precedes it. The motion described by Purkinje was seen following fixation on a stationary pattern. If it consisted of fine, regular, parallel lines then prolonged observation produced an aftereffect of motion perpendicular to the lines in the adapting pattern. This aftereffect has been rediscovered many times since and it is now called the complementary aftereffect. The shimmering movement is also visible during stimulation, particularly with radiating patterns with blank concentric rings in them. Such stimuli have been used in brain imaging studies to determine the cortical location that is most active during the perceived motion.

Purkinje is also known for describing the reflections of external light sources that can be seen in the eye. There are four of these, and they derive from the front and rear surfaces of the cornea and of the lens, respectively. These reflected images can be used to track eye movements, since they displace with respect to the orbit when the eye does. Since the cornea and lens are separated in depth, the first (front surface of the cornea) and fourth (rear surface of the lens) images move relative to one another, and this provides a sensitive method of measurement, commonly known as a dual Purkinje eye-tracker.

Hallucinations

Perhaps the most commonly accepted hallucinations are those that are experienced by some individuals but not corroborated by others – they are beyond common sensory experience. This definition has the advantage of excluding illusions that can be experienced by everyone. For example, what we now call afterimages can be seen following observation of a very intense light. However, there are still problems with this simple definition, and they involve individual differences in experience. Colour deficiencies can be used as an illustration. Differences in the ability to discriminate the range of colours were described remarkably late in the history of vision and it was restricted to a small number of individuals. It was not considered to be a hallucination because a body of data and theory existed regarding colour vision, and colour deficiency could be accommodated within it.

Many hallucinations are a consequence of other changes taking place in the brain. One that is associated with the ageing brain was described by the Swiss naturalist Charles Bonnet (1720–1793) and the syndrome now bears his name. His grandfather “sees from time to time before him, faces of men and women, of

birds, of coaches, of houses etc. … But what it is very important to notice is that this old man does not take, as visionaries do, his visions for reality". That is, he would report seeing objects with clarity while at the same time being aware that they were not actually present. The Charles Bonnet syndrome is associated with vision loss in otherwise healthy individuals; it is still relatively rare, affecting perhaps around 10 per cent of the aged population. With the increase in life expectancy conditions like this are likely to become more common.

Variations in vision

We have most difficulty in perceiving, and therefore in acting efficiently, when the information available to the senses is either greatly restricted, or provided by an environment with which we have no natural experience. Examples of the former are given by the gross errors of judgement that occur when attempting to drive a car at speed in fog or at night, and of the latter when people try to perceive correctly in outer space or under water. In these cases we require special training and experience in order to perceive the 'where', 'when' and 'what' of objects, and without it we are liable to experience distortions of space. There are also instances in which the mechanisms that underpin perception are not functioning appropriately, either as a consequence of peripheral problems (as in the case of most colour deficits) or central ones (as for visual agnosias). Even when the apparatus for vision is malfunctioning it can be very difficult to discover this. Colour deficiency is a case in point: the use of colour names for particular objects in our environment (like green grass) has delayed the appreciation that not all people have the same colour experiences.

These points can be amplified by examining a developmental condition that affects the alignment and movements of the eyes; it is called strabismus or more commonly squint. Distortions of the eyes were recorded by Egyptian physicians over 3,000 years ago, but its reported association with problems of binocular vision are more recent. When the eyes deviate, double vision can ensue. That is, an object is seen in two locations rather than a single one. This is easy to demonstrate by pressing the side of one eye while looking at an object: it will briefly appear double. Our eye muscles work in such exquisite harmony that we are rarely aware of the world appearing double. Strabismus usually occurs early in life and if it is not corrected, either by exercises or surgery, normal binocular function will be lost. Corrections for the deviation were advocated over 1,000 years ago and masks, like those shown in Figure 2.31, were introduced over 400 years ago.

Strabismus can have early or late onset, and the visual consequences are greatly dependent upon this. For late onset (after about 5 years of age) there is initially double vision, but gradually one eye is used in preference to the other to achieve singleness of vision. The individuals become essentially one-eyed, even though the other remains open. That is, the input from one eye suppresses the processing of that from the other and vision in the deviating eye becomes blurred. This is called amblyopia, and it cannot be corrected by any optical means. Thus,

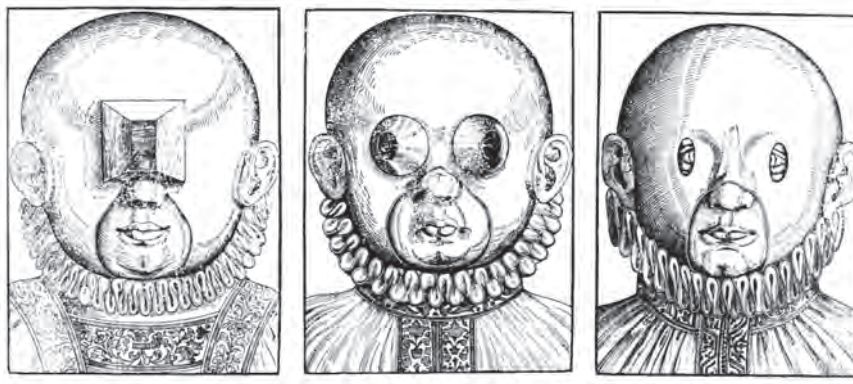


Figure 2.31 Sixteenth-century masks for correcting squints. It was recognised that the earlier the masks were applied the better the chances of correction.

the poor acuity is not due to the optics of the eye, but to the ways the neural signals are processed in the visual system. The misalignment of the eyes is very evident to anyone looking at a person so affected, and is probably one of the reasons why medical corrections for the condition have long been sought. If onset of strabismus is early (in the first few years of life) then corrections or surgery to bring the eyes into alignment can leave binocular vision intact. The differences between early and late onset strabismus suggests that there is a critical period in the development of human binocularity, and this accords with animal studies. If strabismus is induced in monkeys soon after birth then the cells in the visual cortex that normally respond to stimulation of both eyes do not develop. Inducing strabismus in older monkeys does not affect the complement of binocular cells in the visual cortex.

Squinting was not only a medical condition that warranted treatment but it provided a naturally occurring example of retinal non-correspondence to which theories could be addressed. In this way, the study of strabismus not only aided understanding of this condition, but also promoted further research into eye movements themselves and generated new insights into the normal operation of the eyes.

Colour deficits

Colour blindness is a very rare condition, and yet the term is both widely known and widely applied. Very few instances of blindness to all colours (known as rod monochromacy, a very disabling condition associated with albinism) have been encountered, unlike the numerous cases of colour defects and anomalies, in which one colour is either not discriminated from others or is not seen as intensely as for the majority of people. It is remarkable that colour blindness came so late on the scientific scene. It must have been present in human societies for millennia, and yet the first descriptions of it go back barely four centuries. The likely reason

Figure 2.32 Left, **Daltonism** John Dalton (1766–1844) described his own colour defects, but is best known for his advocacy of a theory of atoms. Right, **Light historian** Joseph Priestley (1733–1804) wrote a history of light and colours in 1772, and described a family with colour anomalies.

for this is that colour names can mask differences in colour vision. Nonetheless, when colour descriptions were made without the benefits of object knowledge, differences were recorded. A particularly significant instance involved the chemist John Dalton (Figure 2.32). His attention was drawn to the peculiarities of his own colour vision in the context of botanical specimens viewed by day and by candle light; the geranium flower "was pink, but it appeared to me almost an exact skyblue by day; in candlelight, however, it was astonishingly changed, not having then any blue in it, but being what I called red, a colour which forms a striking contrast to blue". When he asked his brother to look at the flower under similar circumstances, the description was like his own. His friends, on the other hand, remarked that there was little colour change to be seen. This startling discovery of his (and his brother's) defective colour vision started his more systematic investigations of the phenomenon. It is noteworthy that this 'deficiency' resulted in Dalton and his brother seeing more differences in the flower's colour than others did, so we should be cautious about describing it as just a loss of discrimination. Dalton's account had a great impact on the scientific community, presumably because of the high esteem in which he was held for his atomic theory of matter. Indeed, such was the interest that the condition itself became called *Daltonism* in the early nineteenth century.

Dalton's account of his colour anomaly (delivered at the Manchester Literary and Philosophical Society in 1794) concluded: "It is remarkable that I have not heard of one female subject to this peculiarity". There are now known to be marked differences in the incidence of colour defective vision in females and males: only about 0.5 per cent of females have colour anomalies whereas they occur in 8 per cent of males. Dalton believed that his inability to distinguish scarlet from green or pink from blue was because the vitreous humour of his eyes was tinged with blue, thereby absorbing the longer wavelengths of the spectrum. When the hypothesis was tested after his death (on his written

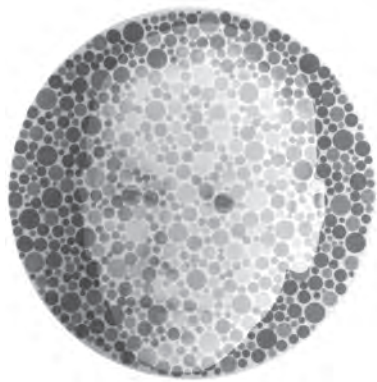


Figure 2.33 Colour blindness Shinobu Ishihara (1879–1963) devised colour plates for detecting individuals who make red–green confusions. He is shown in an achromatic example of one of his colour plates. Instead of a numeral or a line being visible his portrait is defined by slight brightness differences throughout the pattern of dots.

instructions), the humours were found to be colourless. Two centuries after Dalton's description of his own colour defect, retinal pigment from one of his preserved eyes was examined and it was found to be lacking the photopigment that absorbs middle wavelength light. In fact a contemporary of Dalton's, Joseph Priestley (Figure 2.32), had earlier described the colour anomalies of a whole family: the father was colour defective, unlike the mother, but her brother had abnormal colour vision; one daughter was normal, a son and a second daughter were colour defective; the second daughter bore two colour-defective sons and a colour-normal daughter, while the son had a boy and girl who were colour normal. Thus, the essential clues to the hereditary nature of certain colour deficiencies were available in the eighteenth century. The remarkable aspect of this report is that the brother of the mother was colour defective, too, rendering the occurrence of a defective daughter comprehensible in terms of the sex-linked inheritance of recessive genes.

Colour deficits have had an important bearing on theories of colour vision. There has been much debate on the nature and number of primary colours. Normal colour vision is trichromatic. That is, the range of colours that we experience can be matched by appropriate mixtures of just three primaries – corresponding to red, green, and blue parts of the spectrum. Many people with colour defects only require mixtures of two primaries to match any other wavelength presented to them. They are called dichromats, and they are subdivided further into protanopes, deuteranopes, and tritanopes because they are missing the first (red), second (green), or third (blue) detectors, respectively. Dalton was a deuteranope and this is the most common form of colour deficiency. Tritanopes are much rarer than the other dichromats, but monochromats, who can match any wavelength of light with any other, are rarer still. The largest group of colour defective individuals are anomalous trichromats: they can match any wavelength with three others, but the proportions are different from the normal trichromats.

With growing understanding of the variations in colour vision that exist in the population and with a better means of classifying the colour deficits, attention was directed to discovering methods by which they could be tested. One of the earliest, and still used, for distinguishing between those who confuse red and green was devised by Ishihara (Figure 2.33) in 1917. The test consists of colour plates which are called pseudoisochromatic; that is, they have arrays of dots differing markedly in hue (being mainly reddish and greenish) but only slightly in brightness. Within the arrays of dots, numerals or paths in the patterns that are visible to those with normal colour vision remain invisible to those who have difficulties in detecting colours in the red and green part of the spectrum.

Reference notes

One of the features of this book, which is not shared by other introductory textbooks on perception, is its emphasis on the ways in which our ideas about the senses have developed over time. This will be evident from the first chapter but it is amplified in the topics covered here. Indeed, much of the historical interest in vision has concerned its abnormal function rather than its normal, reliable delivery of the state of the external world. Goethe examined a wide range of colour phenomena and his quotation was taken from a translation of *Farbenlehre/ Theory of colours* (1810/1970). Histories of psychology must engage with the history of vision, but not many books are devoted to the latter. Boring's *Sensation and perception in the history of experimental psychology* (1942) stands out and continues to provide a purview of research from the seventeenth to the mid-twentieth century. One of the authors has written *A natural history of vision* which covers the observational era of vision research up to the middle of the nineteenth century (Wade, 1998a), that is, the period prior to the widespread study of visual phenomena in the laboratory. The sources of many of the conventional portraits can be found in this book. The 'perceptual portraits' have been produced by one of us (NW) and in some cases coloured versions of those shown in black and white can be found at http://neuroportraits.eu/ .

Much of the debate during these centuries concerned the nature of light itself. The principal Greek theorists of light were Euclid and Ptolemy, and their texts on optics are available in English. In fact, they were both translated remarkably late, considering their importance: Euclid's *Optics* appeared in English before that of Ptolemy (see Burton, 1945; Smith, 1996). Light and sight were conflated in a variety of ways by Greek thinkers, and their ideas were transmitted and extended by Arabic writers like Ibn al-Haytham, to be reabsorbed into European thought from the thirteenth century onwards to form the mediaeval *Perspectiva*. These historical courses have been charted by Lindberg (1976) and by Russell (2010). Sabra (1989) translated Ibn al-Haytham's optics from Arabic into English. Greek theories were summarised by Stratton (1917), and the Alcmaeon's description of 'fire in the eye' is taken from his book. *The fire within the eye* was the title for Park's (1997) book on the history of light. Kepler was a major figure in directing attention to the eye as an optical instrument. The quotation regarding natural philosophers

was taken from Crombie's (1964) partial translation of Kepler's book on optics; the whole book has been translated by Donahue (2000). Kepler's book was published a century before Newton's *Opticks* of 1704. The study of astronomy has had an enormous impact on our understanding of vision, and some aspects of this are outlined by Wade and Piccolino (2009). The relationship between Grimaldi's comparisons of sound and light waves and Hooke's experiments is discussed in Wade (2005b).

For Plato and Aristotle it was not light but colour that was the principal source of interest in vision. They distinguished between light and colour, considering that light had its ultimate source in the sun, but colour was a property of objects themselves. Thomas Young's life and work have been engagingly presented by Robinson (2006); Young is described as the last man who knew everything! What he did know was presented in his lectures to the Royal Institution, and these have been reprinted (Young, 2002); all Young's articles on vision and optics are available in this book. The changing status of the concept of colour from Aristotle to Newton is clearly detailed by Kemp (1990) in the context of art and science. This applies to colour mixing as well as the representation of space in art, and the emergence of linear perspective as a dominant style in the early fourteenth century. Gombrich has written eloquently about the interplay of vision and art, and it is his book *Art and illusion* (1960) from which the quotation concerning the riddle of style is taken. Magritte asked many searching questions regarding representation, and graphical manipulations of his pipe tune can be found in Wade (1990), from which Figure 2.16 was taken. Further graphical variants on this theme can be found in Wade (2011). Figure 2.17, depicting the principles of depiction following the rules of linear perspective, is from Kirby (1755) as is Figure 2.18. Depth, or the lack of it, has been a constant theme in painting, and Leonardo's quotation regarding the relief in paintings is taken from the first translation into English of his collection of notes on painting (1721). Modern editions of the *Treatise* are available, and that edited by Kemp (2001) is recommended. Leonardo made comparisons between the eye and a camera obscura, as did others later. A history of this remarkable and simple instrument can be found in Hammond's (1981) book. Mannoni's (2000) delightful *The great art of light and shadow* uses the title of a seventeenth-century book concerned with optical instruments. Mannoni does not confine his account to the camera obscura but examines a wide range of other visual toys that played a part in the origins of cinema.

Practical concerns of visual development have concentrated on ageing but theoretical issues have addressed the young. However, rather than examining vision in the young, thought experiments concerning the restoration of sight captured the attention of philosophers. The celebrated question posed by William Molyneux to John Locke was printed in the second edition of the latter's *Essay concerning human understanding* (1694), and the quotation is taken from p. 67. Although Locke was an empiricist philosopher an empirical (experimental) answer to Molyneux's question was not envisaged. Subsequently, there were many attempts to address the question through cataract surgery and these have been assessed by Morgan (1977) who also examined the theoretical issues in detail.

The recent case of Mike May, whose recovery has been charted with the armoury of modern psychophysics and imaging methods, has been described by Fine *et al.* (2003). Despite the insights derived from contemporary investigations into the question posed by Molyneux, the answer he proposed remains questionable (see Wade and Gregory, 2006a, b).

Lloyd Morgan issued his cannon in 1894 in *An introduction to comparative psychology* which essentially defined the area that was to adopt that name, and the quotations are taken from it. The detail applied to observing and interpreting the behaviour of his fox-terrier contrasted starkly with that applied to the young of his own species. The neglect of research on infant vision is evident in the quotation from William James (1890, p. 488) concerning the "great blooming, buzzing confusion". The ingenuity that has been applied to tapping the visual capabilities of newborn and young infants can be found in the books by Atkinson (2002) and Daw (2006). They both examine the influence of early visual stimulation on brain development, as well as the effects of strabismus (squint) on the development of binocularity.

Optics has played an important part in the study of visual development, with the initial emphasis being on changes taking place in the ageing eye. These, and many other late developmental changes in vision are covered in *The aging eye* (2001) compiled by the Harvard Medical School. Understanding the reduction in visual acuity with age was advanced greatly in the nineteenth century, and Figure 2.24 is taken from the most influential book of that era by Donders (1864). Donders introduced the term 'visual acuity' for what had previously been described as the 'minimum visible or separable', and he was quick to apply the new method for assessing acuity in his own studies of declining acuity with age. An account of the historical developments concerning optical imaging and vision can be read in Wade (2007b).

The topic of illusions is a huge one and, as has been alluded to above, it has occupied a disproportionate amount of the research on vision. There are many books on illusions (Gregory, 2009; Ninio, 2001, 2011; Robinson, 1998; Vicario. 2011; Wade, 1982, 2005a) and numerous websites devoted to displaying them. An invaluable source of information on illusions can be found in the *Oxford compendium of visual illusions* (Shapiro and Todorovic, 2013). Studies of illusions that can be seen in the natural environment, like the moon illusion, have been surveyed by Ross and Plug (2002). The motion aftereffect continues to excite experimental interest, and accounts of research on it can be found in the book by Mather, Verstraten and Anstis (1998). It was initially called the waterfall illusion, and the quotation is taken from the original report by Addams (1834, p. 373) and the whole report has been reprinted in Swanston and Wade (1994). Purkinje also wrote about the motion aftereffect, as well as a host of other subjective visual phenomena. Most of these were published in two books that essentially defined the area; the first of these has been translated into English, with a commentary on Purkinje's work, by Wade and Brožek (2001). Purkinje was writing in the 1820s, before the term 'geometrical optical illusion' designated certain two-dimensional figures that induced small distortions of size or orientation. Such illusions have

been catalogued by Robinson (1998) and by Vicario (2008, 2011). Purkinje was working on subjective visual phenomena at the same time as Johannes Müller, whose contributions in this area have been assayed by Berrios (2005). A modern treatment of hallucinations can be found in Aleman and Larøi's (2008) book. They examine the neurological underpinnings of hallucinations, which also provide the focus for studies of agnosias. Bonnet's description of his grandfather's hallucinations is taken from his original account in 1760, as translated in Whitaker and Turgeon (2008, p. 198).

Colour anomalies very rarely involve the absence of cone receptors, as is the case with rod monochromats. There are around three in every million of the normal population; on the Pacific island Sacks visited (Pingelap in Micronesia) the proportion was more than five in 100 (Sacks, 1998). Most common cases of colour anomaly involve variations in the number or sensitivity of cone receptors. This was the case for Dalton, and the quotation regarding his colour anomaly is from his published report in 1798. The photopigments in Dalton's preserved eye were examined by Hunt *et al.* (1995). An account of Priestley's work on vision can be found in Wade (2004a). Colour deficiencies and the genetic factors that are involved in it are clearly presented in Valberg's (2005) book *Light vision colour* and in Mollon, Pokorny, and Knoblauch (2003). Tests of colour vision were introduced increasingly from the late nineteenth century, when the classes of colour deficiencies were more clearly defined. Ishihara's test is still widely used, and the colourless portrait of him is taken from Wade (1992a).

Unlike colour deficiency, studies of and corrections for strabismus have a long history. This is largely because disturbances of binocular single vision (double vision or diplopia) are not only evident perceptually but also they can have critical impact on visual-motor coordination. The masks shown in Figure 2.31 for assisting in aligning the eyes are from Bartisch (1583). Binocular single vision also directs us to the close relationship between the extraocular muscles and vision. The importance of binocular alignment during early development is emphasised by both Atkinson (2002) and Daw (2006), as well as the critical periods during which corresponding stimulation is required for later normal binocular vision. Modern accounts of the development and disorders of binocular vision can be found in McCoun and Reeves (2010).

3 Theoretical approaches to vision

Perceptual experience is subjective. Each one of us is able to reflect upon the nature of that experience, and to describe it to others. While the experience is subjective, and inaccessible to others, the descriptions of what we perceive can be shared with others. Generally speaking, the descriptions given by different people of their perceptual experiences are in remarkably close correspondence; so close, in fact, that many people equate their perception of the world with the way it is described in traditional physics (or what we might loosely call reality). This has been the topic of much philosophical debate and the equation of perception with external reality is often referred to as naive realism. The close correspondence between perception of the world and other descriptions of it (e.g., as in terms of physical measurement) retarded the analysis of perception in general. The aspects that did demand scrutiny were those in which perceptual and physical descriptions did not match, or in which some disease, injury or intervention influenced perception; that is, when perception is no longer veridical or equated with physical measurements.

We can cite two examples of departures from perceptual veridicality described by Aristotle – afterimages and aftereffects. In one set of observations he directed his eyes briefly at the sun and noticed a rapid sequence of colours when he looked away, followed by a dark disc. It was obviously of interest to Aristotle because there was no visibly coloured object that corresponded to the briefly perceived colours. You should not try to repeat this observation, as it could damage the retina. There is an even more remarkable anecdotal account of such folly in the history of afterimages: a friend of the chemist Robert Boyle (1627–1691) looked at the sun through a telescope, and reported that he could see its afterimage 11 years later! It seems more likely that he was not seeing at all with the part of the retina that had been exposed to the concentrated sunlight because the retinal cells would have been destroyed producing a scotoma. Another phenomenon described by Aristotle was the motion aftereffect. He looked for some time at the stones at the bottom of river through the rapidly flowing water, then he directed his gaze to the stones on the bank of the river; these appeared to be moving, too. That is, here was a situation in which the same objects, the stones on the river bank, appeared to be stationary at one time and moving at another. Have the stones changed their location or has their perception been modified? This clearly poses a problem for naive realists,

who believe that reality is equated with our perceptual experience. The movement aftereffect remains a phenomenon of interest to perceptual psychologists, and it can be elicited with a wide range of moving stimuli, as will be discussed in Chapter 7.

Early theories of vision

Philosophy has played a central role in the study of perception because the senses and their functions have been of focal importance to philosophy. Most of the basic ideas were initially expounded by Greek thinkers, and they have been elaborated upon by more modern philosophers. Moreover, it was essentially the unusual aspects of visual experience that elicited most early interest. Of these, three were of particular significance and they influenced theories of light as well as sight; they were discussed briefly in Chapter 2. The first concerned the experience of light following pressure or a blow to the eye; the second related to the visibility of a reflected image in the eye. The idea of light being emitted from the eye was founded on the first of these, and the notion of an image being carried back to the eye was the source of the second. A third feature of sight, which distinguished it from the other senses, was that the experience could be terminated by closing the eyelids during daytime.

Plato distrusted the senses because the evidence they furnished about external objects could change; moreover, the objects themselves could change, as in the process of growth. Accordingly, he believed that the world of appearances was one of illusion, as opposed to the world of thought in which ideal forms existed. The forms reflected the universal qualities of objects rather than the particular features which can be sensed. The abstract forms could be investigated by reasoning rather than observation, and this resulted in a preference for rational rather than empirical enquiry. Plato's position demonstrates the influence that language has had on philosophical thought: particular members of a category that are given a single name (e.g., horse) do not reflect their universal characteristics. These ideal forms are permanent and inaccessible to perception because the senses are concerned with particulars rather than universals. Plato distinguished between the body and the soul: the body was part of the material world whereas the soul was immaterial. He likened the rational soul to a charioteer steering the competing horses of emotion and appetite; the rational soul was considered to be morally superior to the others and should guide their actions. These distinctions were to have considerable significance because they later permeated both philosophy and Christian theology. Mind–body dualism was at the heart of Descartes' philosophy as well as a constant current in Christian theology. The latter also placed great emphasis on the moral superiority of reason over irrational feelings and passions.

Philosophy

Aristotle adopted more naturalistic explanations of phenomena which did not denigrate the senses. He criticised Plato's theory of ideal forms, arguing that the features that distinguish a horse, say, do not have an existence independently of

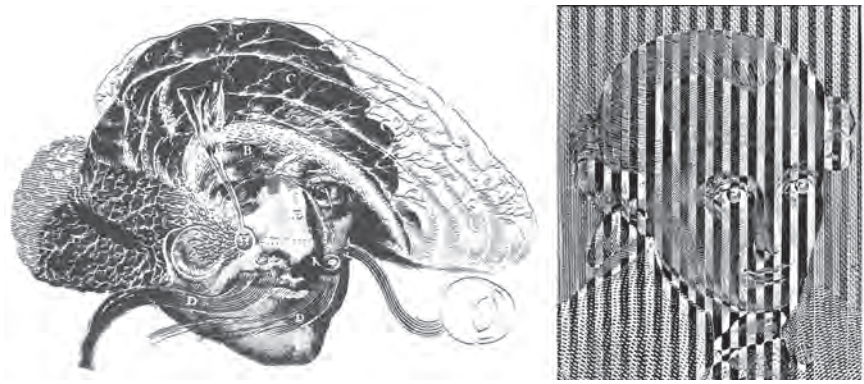


Figure 3.1 Left, **The dualism of Descartes** The portrait of René Descartes, after the famous painting by Frans Hals, is combined with an engraving of the brain taken from Descartes' *Treatise of man* (*De Homine*). He wrote: "Conceive surface AA facing cavities EE to be a rather dense and compact net or mesh, all of whose links are so many little conduits which the spirits can enter and which, always facing toward gland H [the pineal] whence the spirits emanate." Descartes considered that movements of the pineal gland (controlled by the mind) redirected the animal spirits (carried along the nerves) to move the muscles. Right, **Apperception** Immanuel Kant (1724–1804) developed a transcendental theory of mind which drew upon both rationalism and empiricism without being allied to either. Kant replaced Descartes' self or ego as the ultimate reality with "the transcendental unity of apperception". Apperception referred to the active ordering of our experience to achieve meaning, and it is applied in recognising Kant's portrait: neighbouring vertical bands are in opposite contrasts, but despite this the figure is seen as representing a face against a patterned background.

horses; these distinctive features could best be studied by examining actual horses rather than their ideal forms. Therefore Aristotle preferred an empirical approach to a rational one. He is often considered to be the first psychologist because of his emphasis on observation and because he tried to order phenomena in a systematic manner.

Rationalism

The contrast between the philosophies of Plato and Aristotle can be thought of as that between rationalism and empiricism, and both approaches have been applied to the analysis of perception. The differences between them were brought into sharper focus after the scientific renaissance in the seventeenth century, when a wider range of perceptual phenomena was scrutinised. Descartes (Figure 3.1) applied mechanistic interpretations to bodily processes while maintaining that the mind was immaterial, thus retaining the Platonic distinction between body and soul. Communion between mind and body was achieved via structures in the brain, particularly the unpaired pineal body.

Descartes' mechanistic approach to the senses clarified many issues in perception, but he had the thorny problem of accounting for the interaction of

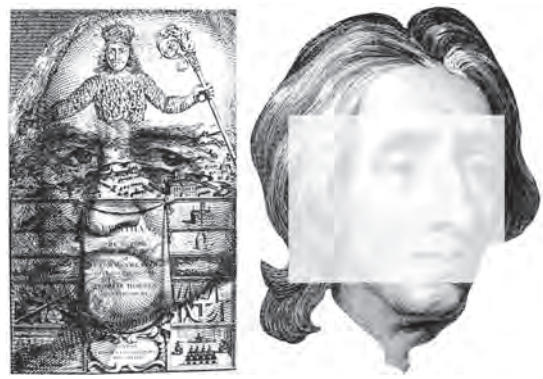


Figure 3.2 Left, **Leviathan** Thomas Hobbes (1588–1679) is portrayed as part of the Leviathan, the macrocosm of the state that is inseparable from the microcosm of man: "the multitude so united in one person, is called a Commonwealth ... This is the generation of that great Leviathan". Right, ***Tabula rasa*** John Locke (1632–1704) rejected one of Descartes' principal concepts, that of innate ideas, initiating the continuing debate between nativists and empiricists. Instead of accepting that some ideas, such as unity or God, are inborn, Locke proposed that all knowledge derives from experience. The portrait of Locke represents the importance of the senses in his philosophy: the eyes, ears, nose, and mouth have been covered by a blank, transparent sheet of paper (*tabula rasa*).

the rational mind with the mechanistic body. This was a task attempted later by Immanuel Kant (Figure 3.1). Kant did not deny that all knowledge begins with experience, but he did not believe that it all arises out of experience either. He considered that certain aspects of knowledge were innate, most particularly the ideas of space and time. That is, Kant suggested that the individual is born with the ability to organise experience in both space and time. Perception is then an active organising rather than a passive receptive process. Kant's influence on Continental philosophy was vast, but it also had numerous repercussions in related disciplines like physiology and psychology. The distinction between innate and learned processes in perception became enshrined in nativist and empiricist philosophies, respectively. The nativists believed that we are born with the ability to perceive space, whereas the empiricists argued that we have no such knowledge of the world at birth, but we need to learn to see the spatial attributes like size, shape and distance. For Kant our conscious, phenomenal world is a cognitive construction. He made a distinction between the world of things and that of appearances, and was pessimistic about whether the latter (and hence psychology) was open to scientific enquiry. That is, he did not consider that the inner world was open to precise measurement, and therefore its study could not be classified as a science.

Empiricism

The foundations for empiricist philosophy were laid by Thomas Hobbes (Figure 3.2) in the seventeenth century, although the basic ideas about learning

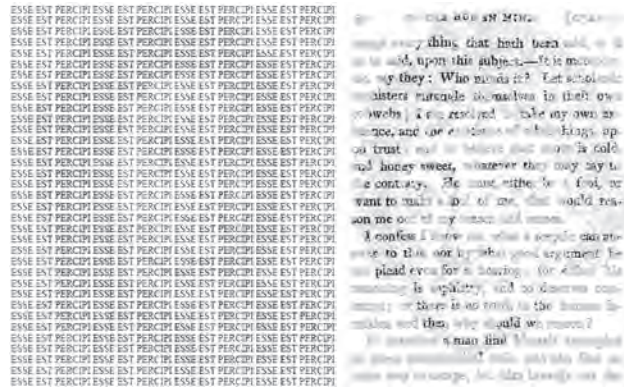


Figure 3.3 Left, **Berkeley's esse** The portrait of George Berkeley (1685–1753) exists if it can be perceived. Right, **Reiding the mind** Thomas Reid (1710–1796) is shown in text from his book *An inquiry into the human mind, on the principles of common sense*; the text describes his common sense approach to perception while pouring scorn on the philosophies of Berkeley and Hume.

and perception have a much longer history. Hobbes placed the senses as the source of all knowledge: "there is no conception in a mans mind, which hath not at first, totally, or by parts, been begotten upon the organs of Sense". In this regard, his empiricist ideas were in direct conflict with Descartes' dualism. Hobbes also presented a fledgling theory of association. The "Trayne of Thoughts" were a consequence of their coherence (or simultaneous occurrence) and could be random or regulated; the latter were guided by desire and habit, and associations could be strengthened by frequent repetition.

Modern empiricist philosophy was expounded by John Locke (Figure 3.2) at the end of the seventeenth century. In his *Essay concerning human understanding*, published in 1690, he wrote: "Let us suppose the Mind to be, as we say, white Paper, void of all Characters, without any *Ideas*; How comes it to be furnished? Whence comes it by that vast store, which the busie and boundless Fancy of Man has painted on it with an almost endless variety? Whence has it all the materials of Reason and Knowledge? To this I answer in one word, from *Experience*: In that all our Knowledge is founded, and from that it ultimately derives it self". For Locke the mental element is the idea, which is based upon sensory experience. Ideas could be simple (like whiteness) or compound (like snow), and compound ideas are made up from associations between simple ones, by a process like 'mental chemistry'. Similar associative links can account for our ability to generalise across stimuli: for instance, to form a general idea of a triangle from many different specific instances. Thus, Locke was an empiricist and an associationist: knowledge derives from the senses and we learn to perceive the objects in the world by association.

Locke charted the course for empiricism, but many of the details were provided by later philosophers, two of whom will be mentioned briefly here. George

Berkeley (Figure 3.3) argued that we learn to perceive the dimensions of space by associating muscular sensations with those of vision. In order to perceive distance visually we learn the relationship between the visual stimulation and the states of the muscles controlling the eyes. The muscular and touch systems were considered to provide direct and undistorted spatial information that could be used to teach vision the dimensions of space. Berkeley refined the empiricist philosophy of Locke by arguing that appearances are all: to be is to be perceived – *esse est percipi*. That is, the matter from which materialism is constructed is itself open to question. If all we have are our perceptions, how can we prove the existence of an external world? A problem with this position is that if perceptions are transitory so is existence. Does an object cease to exist when the eyes are closed? Berkeley sought to salvage this slide into solipsism (that nothing other than one's own ideas exists) by arguing that God alone perceived an external reality. Despite this idealist stance, Berkeley made important steps towards understanding how we perceive space, and how the different spatial senses are integrated.

Thomas Reid (Figure 3.3) reacted to Berkeley's idealism by arguing that the evidence of external reality is provided by the common activities of the senses and is supported by common sense intuition. Reid founded the Scottish common-sense school of philosophy, whose ideas were to be influential in the development of psychology in America in the nineteenth century. The school was opposed to associationism, particularly when it was couched in physiological language. Reid also proposed a faculty psychology; faculties were innate properties of the mind which exerted control over habits, or behaviour. His descriptive psychology could be studied by reflection on mental activity, by an analysis of the use of language, and by observations of behaviour. He provided a bridge between the extreme rationalists and empiricists. His belief in the power of reason was tempered by a desire to accumulate evidence empirically. Reid is perhaps best remembered because of the clear distinction he made between sensation and perception. Thus, redness and roundness may be sensations produced by an apple, but its perception includes an appreciation of the object itself. Perceptions also involve projective aspects that are not present in sensations: the apple is perceived as being out there, but the sensations can be internal. Reid's distinction has had far-reaching consequences, and it has persisted to the present; it has pervaded our language and it even defines the categories of our enquiries.

Empiricist philosophy was initially confined to Britain, but its widest influence has probably been through its adoption beyond Britain's shores – particularly by Helmholtz in Germany and Watson in America. Helmholtz carried the torch for philosophical empiricism in a hostile Kantian climate, as did his erstwhile assistant Wilhelm Wundt. However, their brands of empiricism were quite different. Helmholtz developed the notion of unconscious inference to account for characteristics of colour and space perception, and the concept is still active in some theories. Wundt was more ambitious and applied empiricist and associationist ideas to account for consciousness itself.

Nineteenth-century influences

Discussions of space and time continued to be grist to the philosophers' mill throughout the nineteenth century, but they were afforded instrumental and experimental assistance that extended their scope. It can be argued that the evolution of psychology as an independent discipline was in large part a consequence of addressing philosophical questions concerning the perception of space and time by recourse to experiment rather than exposition.

Phenomenology

Perceptual experience has always been described in words, when possible, but this has not been the only way of assessing it. Language reflects the nuances that can be applied to the richness of perception, and it has been considered by some, like Goethe, to be the most appropriate vehicle for conveying experience. Goethe, in line with many Romantic philosophers, rejected the experimental approach to the study of nature because it was too constrained. In its place he proposed the astute and intuitive observation of natural phenomena, setting in train the method of phenomenology. This is best seen in his 1810 book *Zur Farbenlehre*, which contrasted his observational approach to colour with what he considered to be the physicalism of Newton. The purity of white light was taken to be fundamental and indivisible, rather than white being a mixture of different coloured lights. Goethe was an astute observer and described many colour phenomena. He distinguished between what he called physiological colours (the experience of colour) and physical colours produced by optical refraction. His interpretations emphasised the subjective dimension in perception, and he reflected the ideals of the Romantic movement in art and science.

Phenomenology was given a more methodological twist by Purkinje (Figure 3.4), whose interests in vision were stimulated by reading Goethe's analysis of colour. Purkinje was encouraged in his researches by Goethe because of his use of the phenomenological description. As a medical student Purkinje investigated subjective visual phenomena in part because he did not have access to any physiological apparatus, but also because he believed that visual illusions revealed visual truths. Most of his experimental research in both physiology and histology was conducted in Germany, but at the age of sixty-three he was called to the chair of physiology in Prague. He was followed in that chair by Ewald Hering (Figure 3.5), who also embraced phenomenology.

Nativism versus empiricism

The contributions made by Helmholtz (Figure 3.5) to visual science are legion, but his most lasting impact was his theory of perception. For Helmholtz the brain only had indirect access to the external world, via the senses, and it could only process messages in the language of nerve impulses. This realisation made any equation of the retinal image with perception unnecessary, and it removed a problem that

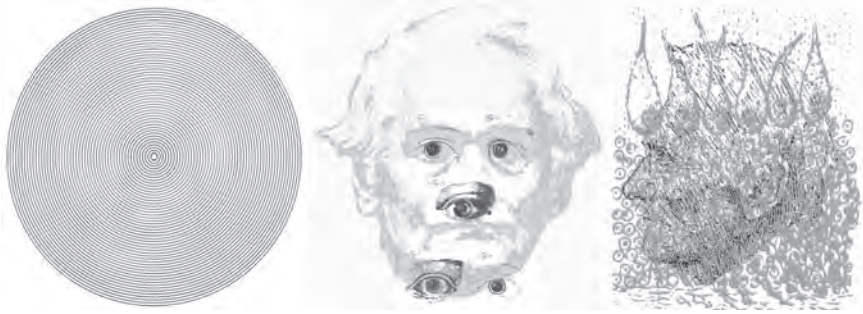


Figure 3.4 **Images of Purkinje** Jan Evangelista Purkinje (1787–1869) made many contributions to science in addition to his support for phenomenology. His portrait is shown three times. On the left, his face is partially hidden in a pattern of concentric circles (like that in Figure 2.30). In the centre, a full face portrait is combined with his illustration of the reflections of a candle flame in the eye (now known as Purkinje images). On the right, he is embedded in his representation of the cells in the cerebellum that are now called Purkinje cells.

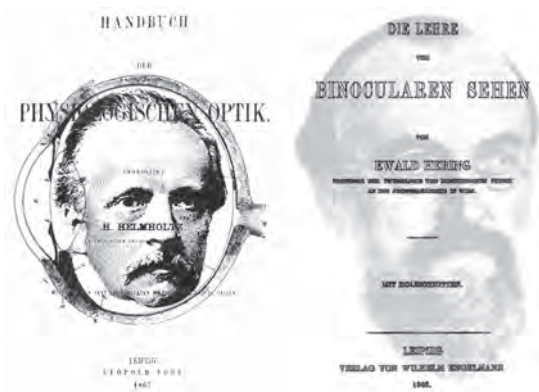



Figure 3.5 ***Empiricist and nativist*** Left, Helmholtz is shown in his diagram of the eye and with the title page of his *Handbuch*, the complete three volumes of which were published in 1867. In it, the empiricist account of vision was boldly stated. Right, Karl Ewald Konstantin Hering (1834–1918) published his theory of binocular vision in the following year. It contrasted his nativist theory of vision with the empiricism of Helmholtz.

had frequently been raised earlier, and was to return later: if the image on the retina is inverted and left–right reversed, why is our perception not so? Helmholtz argued that this only created a problem if there was a picture in the retina that required further perception. If all that is available are nerve impulses then the brain can analyse them and make the appropriate inferences independently of the orientation of stimulation with respect to the retina.

Helmholtz acknowledged that little he wrote on theories of vision was novel, but he marshalled the arguments over a wider range of phenomena than others had done before. By adopting a starkly empiricist interpretation of perception, and by contrasting it so sharply with nativism, he reopened a debate that has reverberated

throughout perception ever since. The debate was personified in the conflict between Helmholtz and Hering, and the main battle-grounds were colour vision and stereoscopic depth perception. Helmholtz summarised his position succinctly:

> *The sensations of the senses are tokens for our consciousness, it being left to our intelligence to learn how to comprehend their meaning* … Evidently, any other sensations, not only of sight but of other senses also, produced by a visible object when we move our eyes or our body so as to look at the object from different sides or to touch it, etc., may be learned by experience. The content of all these possible sensations united in a total idea forms our *idea* (*Vorstellung*) of the body; and we call it *perception* when it is reinforced by actual sensations … The only psychic activity required for this purpose is the regularly recurrent association between two ideas which have often been connected before.

Hering was a physiologist whose psychology was in the tradition of Goethe. He represented the phenomenological and nativist position in studying perception. His work in vision concerned space perception, colour vision and contrast phenomena. Like Goethe, Hering stressed the subjective dimension of colour, and he based his opponent-process theory on colour appearances rather than on mixing lights of different wavelengths after the manner of Young, Maxwell, and Helmholtz. He adopted the procedure of presenting coloured papers to observers and asking them to name the colours from which they were mixed. Red, green, blue, and yellow were not said to be mixtures of any other colours, nor were black and white. He also examined simultaneous and successive colour contrast phenomena. Together these led him to propose a theory of colour vision based on three oppositional pairs: red-green, blue-yellow, and white-black. He speculated that there are three retinal pigments that are either built up or broken down by light to yield the six elements. Modern colour theory has shown both Helmholtz and Hering to be correct in principle but wrong in detail: the initial stage involves three cone pigments (not three kinds of fibres) the signals from which combine neurally (not in the action of the pigments) to produce opposing pairs of red-green, blue-yellow and white-black.

In the area of space perception Hering utilised the concept of local sign; each retinal point was considered to have a local sign for height, width, and depth. This conflicted with Helmholtz's emphasis on learning to interpret the retinal signals, largely via information from eye movements. Hering did investigate binocular eye movements and argued that the two eyes move as a single unit: his law of equal innervation states that when one eye moves the other moves with equal amplitude and velocity, either in the same or the opposite direction. Eye movements were also implicated in visual direction: "For any given two corresponding lines of direction, or visual lines, there is in visual space a single visual direction line upon which appears everything which actually lies in the pair of visual lines". The centre of visual direction was called the cyclopean eye. Hering also attracted a large body of able and loyal students who would carry the nativist banner into the twentieth century.

The psychology of the senses led Helmholtz into the metaphysical domain he had assiduously avoided in his physical and physiological endeavours. The philosophical problems remain a matter of constant revision and re-analysis, but many of the issues concerned with the senses that were debated in Helmholtz's day became the topics of experimental enquiry in the then new discipline of psychology.

Psychology

The psychological dimension of vision was exposed as soon as there were phenomena that warranted recording. The experience of light when pressure is applied to the eye is an early example to which reference has already been made. However, its psychological significance was not appreciated because the various stages in the visual process – the physics of light and the physiology of sight – had not been distinguished. Aristotle did describe many visual phenomena, but it was Ptolemy who realised that the spatial vision could not be understood in terms of visual angles. The psychological dimension increased in prominence after the dioptrics of the eye had been adequately analysed. Nonetheless, they were not a part of psychology because there was then no such subject.

Psychology, as an independent discipline, is considered to have been founded in 1879, when Wundt opened his Psychological Institute at Leipzig University. Prior to that psychology was allied principally to philosophy, although perception was often the province of sensory physiologists. Wundt saw the task of his new institute as that of studying conscious experience. What distinguished his approach from the many earlier ones addressing the same issues were the methods employed. Psychology came of age when it developed its own methodology: the problems of consciousness and perception were examined in novel ways using novel instruments, and psychology became an experimental discipline rather than just an observational one. Perception has followed the theoretical fashions of psychology generally, ebbing and flowing as the subjective dimension waned and waxed in importance.

Wundt incorporated precise methods for measuring detection and discrimination in his Psychological Institute, and it was in providing alternatives to phenomenology that psychology gained its independence from other disciplines. Wundt used analytic introspection for studying consciousness, but in the area of perception he also applied the psychophysical methods that had been described a few years earlier by his compatriots Weber and Fechner (see Figures 1.2 and 1.3). As described in Chapter 1, Fechner integrated Weber's fraction to found the discipline of psychophysics. Fechner's insight was that the mental and material worlds could be united mathematically in the domains of sensory and stimulus intensities: "Psychophysics should be understood here as an exact theory of the functionally dependent relations of body and soul or, more generally, of the material and the mental, of the physical and the psychological worlds". Fechner distinguished between an outer and an inner psychophysics; the former was concerned with the sensation and stimulus intensities, and the latter with the

relation between brain process and sensations. He realised that experiments in his day would be confined to outer psychophysics, but these were seen as necessary steps towards understanding inner psychophysics.

Wundt and structuralism

Wundt (Figure 3.6) rejected phenomenology and introduced a technique that is now called analytic introspection – the controlled analysis of mental events. He distinguished between the mediated experience available to the physical sciences and the immediate experience investigated by psychology. His use of introspection to study the latter resulted in the proposal that sensations and feelings were the elements of consciousness. Sensations could be combined to yield perceptions, but for these to influence behaviour they required attention: the voluntary control of attention to focus on aspects of perception was termed apperception. It was the active role played by attention that could rearrange perceptions to form a creative synthesis. Wundt was an empiricist and an associationist interested in the universal aspects of conscious experience; the application of psychology to real-world issues held little appeal for him.

Wundt himself did measure thresholds, but his greatest interest was the study of conscious experience. Introspection is looking inward at mental processes and the method differed from phenomenology in that instead of using familiar object names (like book or page) only terms signifying sensory quality (colour or brightness),

Figure 3.6 Left, **The institution of psychology** Wilhelm Maximilian Wundt (1832–1920) founded the first institute of experimental psychology at Leipzig in 1879. This event is often taken as marking the onset of psychology as an independent discipline. His portrait is presented within an outline of the building that housed the original Institut für experimentelle Psychologie (which was previously the university prison). Right, **Structuralist** Edward Bradford Titchener (1867–1927) followed in the path of his mentor, Wundt, but he placed more emphasis on the elements of consciousness and their combination by association. His portrait is shown in the circles of the size-contrast illusion bearing his name: the central circles are the same size but that surrounded by smaller circles appears larger than the one surrounded by larger ones.

intensity (extension) and direction were permitted. Thus, the description of the book you are reading could take the form of something like "two white surfaces, separated at their centres with one raised with respect to the other; on each surface are black marks, slightly separated from one another horizontally but with larger spaces at irregular intervals". Observers had to undergo extensive training before they were considered to be skilled at analytic introspection. By using this method Wundt believed that he could determine the elements from which perceptions and thoughts were constructed and he was greatly influenced by Locke's ideas about mental chemistry. Wundt was trying to isolate the basic elements and to determine the rules for their combination into more complex perceptions and thought, and this approach was later called structuralism. The basic elements were taken to be the sensory attributes (like quality and extension), and these could be combined to make the molecules of perception; the combination was achieved by a process of association. Perception represented a synthesis or building up of the sensory attributes via learning by association, whereas the method analysed or broke down complex perceptions into their component sensory attributes.

Wundt attracted many graduate students from the United States, where psychology was forging a strong foothold. One of his students, Edward Bradford Titchener (Figure 3.6), was from Britain but carried Wundt's method and theory to America. Titchener employed introspection to uncover the structures of human consciousness, thereby giving structuralism its name. The elements of consciousness were said to be sensations, images, and affections. These elements could be isolated by a method of introspection that excluded the use of object names, because describing the meaning of objects introduced 'stimulus error'. Most of his experimental work was directed at the analysis of sensations, which had the attributes of quality, duration, intensity, extensity, and clearness. Ideas and emotions derived from associations of images and affections, respectively.

Twentieth-century developments

The advances in the experimental psychology of perception throughout the nineteenth century were driven by the invention of novel instruments that could be used within the laboratory. That is, the study of perception moved from the natural environment into the laboratory where the well-tried methods of physics could be applied to the examination of perception. However, towards the end of the twentieth century the armoury of instruments was replaced by a single device – the computer. The implications of this technological revolution were enormous but they did not play a part in the theories that were formulated in response to the impact of Wundt at the beginning of the century.

Many of Wundt's contemporaries adopted his technique but few were able to obtain the same results as Wundt: analytic introspection was not a reliable means for plumbing the processes of perception. By the early twentieth century there was widespread disaffection with the method and its attendant theory, and alternatives were sought. Two major reactions that appeared were Gestalt psychology and behaviourism. In the second decade of the century, John

Figure 3.7 **Behaviourist** John Broadus Watson (1878–1958) proposed a radical behaviourism that sought to predict all behaviour in terms of S-R bonds. He is represented in the context of sequences of S-R bonds, produced with the mechanical precision of a typewriter.

Watson (Figure 3.7) redefined psychology as the study of behaviour, and he turned his back on its short history as the study of conscious experience. He avoided working with human subjects because he considered that introspection was unreliable and an unsuitable method on which to base any science, and so established the rat and the maze as the subjects for psychology. His views were both radical and initially unpopular, but they were propagated with a religious fervour. Watson launched the behaviourist attack on structuralism in 1913. His dissatisfaction was with the method rather than the theory; in fact behaviourist theory was also empiricist and associationist. The method of analytic introspection was rejected because it was subjective. Sensations and perceptions were inferences based upon introspections, and were not open to public scrutiny as would be expected of a science. Watson's manifesto was clearly stated: "Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior. Introspection forms no essential part of its methods, nor is the scientific value of its data dependent upon the readiness with which they lend themselves to interpretation in terms of consciousness".

Watson argued that the only aspects of psychology that could be measured reliably were the stimuli (S) presented to subjects and the responses (R) they made. Hence, behaviourism was often referred to as S-R theory; the organism was likened to a black box about which nothing could be known directly, but only by inference. Watson and the growing band of behaviourists in America distrusted the study of perception generally, because it could evidently take place without any obvious response. When it was studied, it was in the context of discrimination learning, where the emphasis was more on the process of learning than on perception. Those behaviourists interested in human perception tended to measure overt aspects of it like eye movements. Thus, the Gestaltists became the heirs to perceptual research, almost by default.

Gestalt psychology

Gestalt is a German word that can be translated approximately as configuration, but the German term is retained because of the difficulty of capturing its nuances with a single English word. The Gestalt psychologists were in the main stream of Continental philosophy; they were nativists and phenomenologists. They were nativists because they believed that perception was unitary, and reflected an innate organisation within the brain. Perception was to them an organised process that operated according to rules, and it could not be broken down into its constituent parts. They were phenomenologists because they considered that the subtleties of perception could only be recorded adequately by a system as rich as language. The Gestaltists main opposition to Wundt's structuralism was theoretical – they did not accept that unitary perceptions could be analysed into smaller parts. Indeed, the cliché associated with Gestalt psychology is that "the whole is different from the sum of its parts" – thus, the perception of a square is different from the separate effects from its four constituent sides.

Max Wertheimer (Figure 3.8) redefined psychology as the study of configurations or *Gestalten*. He conducted a series of experiments on apparent movement – motion seen between two stationary and separated stimuli when presented in rapid succession; it was called the phi-phenomenon. The inability to distinguish between real and apparent motion was taken as damning any approach that explained perception in terms its successive sensations. Not only was it said that the whole is different from the sum of its parts, but the perception of the whole is prior to that of its parts. Publication of Wertheimer's thesis on the phi phenomenon, in 1912, is taken as the origin of Gestalt psychology; it was principally concerned with perception, and a range of robust phenomena was devised to support its holistic nature. Much of its attraction lay in the power of the perceptual demonstrations.

Kurt Koffka (Figure 3.8) was the second member of the Gestalt triumvirate. He used Gestalt concepts in studies of development and thinking, and he made

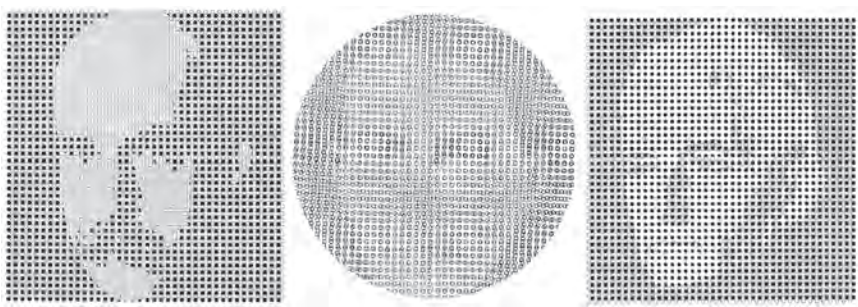


Figure 3.8 **Good figures** The portraits of (left) Max Wertheimer (1880–1943), (centre) Kurt Koffka (1886–1941) and (right) Wolfgang Köhler (1887–1967) are all composed of small circles. It was from combinations of filled and unfilled circles that Wertheimer demonstrated his principles of perceptual organisation in 1923.

American psychologists aware of the new movement in his writings and lectures on Gestalt psychology in the United States. Koffka posed the fundamental question of "Why do things look as they do?" He also emphasised that visual perception is three-dimensional and that our perception is in terms of the object properties (the distal stimulus) rather than those at the receptor surface (the proximal stimulus). The third member, Wolfgang Köhler (Figure 3.8), introduced the concept of field forces operating in both perception and in its underlying neurophysiology: brain processes were considered to be isomorphic (having the same form) with the percept, so that principles of brain function could be inferred from perceptual phenomena. He developed a speculative neurophysiology based mainly on the principles of perceptual grouping and on his experiments with figural aftereffects. His speculations probably did more to hasten the demise of Gestalt theory than any other factor: neurophysiologists failed to find any evidence for such fields of electrical activity in the brain, and so tended to dismiss Gestalt theory in general rather than Köhler's unsuccessful attempt at neuroreductionism in particular.

The Gestalt psychologists formulated some descriptive rules for perceptual organisation and produced a wide range of demonstrations that could be used to support them. The principles were described by Wertheimer in two papers published in 1923; they appeared in the journal *Psychologische Forschung* (later renamed *Psychological Research*) which the Gestalt psychologists founded to propagate their theory. The figures used by Wertheimer consisted mainly of open and closed circles, hence the use of these motifs to carry their portraits in Figure 3.8. The initial and fundamental perceptual process is the separation of a figure from its background, because all the other grouping principles can only operate with segregated figures. Normally, a figure is defined by contours that surround it completely, whereas the ground is larger or lacking a defined boundary, as in Figure 3.9a. Under certain circumstances neither of these conditions is met (see Figure 3.9b), and perceptual instability ensues – first one part and then the other is seen as figure, and this perceptual alternation continues.

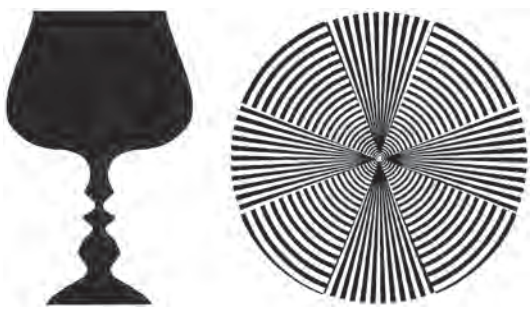


Figure 3.9 (a) A figure, like this black goblet-shape, has closed contours and is smaller than the background against which it is presented. (b) Rubin's crosses. Two crosses, comprised of either radiating lines or concentric arcs, have common boundaries, and fluctuate between being figure and ground, but only one can be seen as figure at one time.

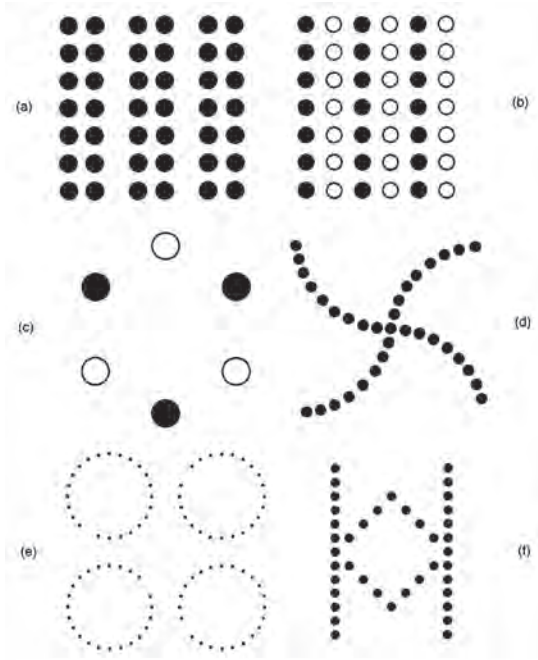


Figure 3.10 Gestalt grouping principles: (a) proximity; (b) similarity; (c) symmetry; (d) good continuation; (e) closure; (f) an example of embedded figures.

In general, pictures of objects are recognised more readily in an upright orientation. This applies particularly to pictures of faces and it has been an aspect of artistic manipulation for centuries. Köhler described the effect clearly: "For this experiment I select a picture, or outline-drawing of an object, which shows a conspicuous change in appearance when it is upside down. This is the case, for instance, with photographs of known or unknown persons. They change so much that what we call facial expression disappears almost entirely in the abnormal orientation". An impression of this can be obtained by looking at any of the portraits in this chapter with the book upside down. It is much more evident when viewing a photograph of a familiar face when it is inverted; not only is it difficult to determine the emotional expression on a face but also identifying the face can be disrupted. This dependence of recognition on orientation can be seen by inverting Figure 3.9a. The black shape that was seen as a goblet is transformed into two swan-necked profiles of white faces against a black background.

Most of the remaining demonstrations of Gestalt grouping principles have clearly segregated figures; they are usually outline drawings, like those shown in Figure 3.10, and these are shown to observers who are asked to describe what they see. For example, Figure 3.10a is said to look like three columns or three pairs of lines. Although many alternative descriptions are possible (for instance, seven rows of dots) they are rarely given. The elements of the configuration are equally

sized dots but these tend to be grouped to form vertical lines; the lines are similar in length and orientation but they differ with regard to their distance or proximity from one another. This was called grouping by proximity.

In Figure 3.10b the dots tend to be described as three columns of black and three columns of white dots. Although all the dots are equally spaced, so that proximity cannot be operating, they are grouped according to their similarity; other things being equal, similar elements within a larger configuration will be related perceptually. The organising principles rarely operate in isolation, more frequently they complement or counteract one another. For example, in Figure 3.10c the dots are typically described as forming two symmetrical triangles, one of white dots and the other of black dots. These are symmetrical figures, and also what the Gestalt psychologists called good figures; simple geometrical shapes, like triangles, squares and circles were considered to be good figures because they could not be reduced perceptually to any simpler components. In Figure 3.10c the principle of symmetry is operating, but not in isolation: it is complemented by the similarity of the dots (black or white), but it is acting against the proximity of the dots.

The dots in Figure 3.10d would be most commonly described as following two intersecting curved lines, rather than alternatives like two V-shapes meeting at their points. This organising principle was referred to as good continuation – the lines are seen as maintaining some continuity of direction and not changing direction abruptly. Even the sequences of dots in Figure 3.10e display goodness of figure, in this case circularity. They also illustrate another Gestalt principle, namely that of closure. In each of the four patterns one dot is missing from the regular sequence, but it is not immediately evident. Any irregularities in good figures tend to be smoothed out perceptually. In Figure 3.10f the various organising principles operate in a way to conceal an aspect of the pattern. This would be described generally as a diamond flanked by two vertical lines, but rarely as a letter W above a letter M. In this instance we are dealing with embedded figures, which are hidden by the grouping rules to yield alternative organisations.

Many more organising principles have been described by Gestalt psychologists, although these are the main ones. Their intention initially was to provide an alternative theory of active, innately organised perception to counter the passive, structuralist views of Wundt and his adherents. The theory was supported by these demonstrations, which drew upon phenomenology. However, it should be noted that the demonstrations themselves were not representative of normal object perception because they were based upon line drawings. That is, the evidence for the principles of organisation is based upon the manner in which two-dimensional pictures are perceived rather than three-dimensional objects. The Gestalt psychologists did conduct more conventional experiments on phenomena like figural aftereffects (see Figure 3.11) and illusions using responses other than verbal descriptions. The results of these experiments were used to suggest a speculative neurophysiology of vision that involved electrical fields in the brain. This strategy was not a wise one because the opponents of Gestalt could attack the physiological speculations far more easily than the robust perceptual demonstrations.

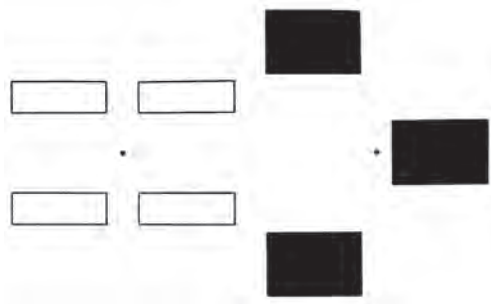


Figure 3.11 A figural aftereffect can be produced by observing these patterns. Initially fixate on the left-hand dot: the left and right pair of outline rectangles will appear to be equally separated from one another. Now fixate on the right-hand dot for about 1 min, after which shift fixation back to the dot on the left. The outlines will no longer look equally separated – the left pair will seem nearer to one another than will the right pair.

Cognitive approaches

In the second quarter of the twentieth century most research in perception was conducted by Gestalt psychologists, initially in Germany and later in America, where it challenged the prevailing climate of behaviourism. At the same time, an alternative approach was being developed in relative isolation in Britain, and it has had a profound effect on the shape of modern perceptual research. Frederic Bartlett (Figure 3.12) examined perception in realistic and dynamic situations and he represents a continuation of the British empiricist tradition with his analysis of perception as a skilled activity. He rejected the application of stimulus-response interpretations of complex tasks (like playing cricket or tennis) because the actions were highly organised and initiated in advance of any contact with the ball. Indeed, the actions were made with respect to the position the ball would be predicted to occupy at some short time in the future. Complex activities of this type indicated that behavioural sequences had to be programmed in advance and coordinated with predictions based on perception. This led Bartlett to a cognitive theory of perception, one in which the division between perception and thought was difficult to draw. Like Helmholtz, Bartlett considered that perception was like problem solving, incorporating processes of inference but also of prediction.

Bartlett rejected associationist models of perception and memory and his cognitive theory provided a middle road between those of the molecular behaviourists and molar Gestaltists. He can be seen as ushering in the cognitive revolution that was eventually to replace behaviourism, although his work was neglected in America until the 1950s. In order to make predictions that involve action we need to have some mental representation of the environment in which the action will take place. This concept of forming a mental model of the world in which we behave was proposed by Craik (see Figure 1.7) and it is one of the ideas that has proved important in the development of both cognitive and computational

theories of vision. The machine metaphor has proved to be particularly attractive to experimental psychologists. Craik was only able to enlist relatively simple machines, but his insight lies at the heart of the cognitive revolution that was to sweep through psychology. His concern with prediction rather than reaction (shared with Bartlett) reflected his dissatisfaction with behaviourism. Perception is considered to be a process in which information regarding aspects of the world is analysed and utilised to plan behaviour. This information-processing approach has become widely accepted as perception can then be considered as a sequence of representations that are initially crude and become increasingly appropriate to the three-dimensional environment.

Information in visual patterns could be quantified, and it resided at the boundaries between areas (contours) and where the contours changed direction abruptly (corners). However, it was the qualitative concept of information processing rather than quantitative information measures that was to have lasting appeal. The perceiver was conceived of as a limited capacity information processor, and the information could be filtered, filed, or reformulated on the basis of stored events. Donald Broadbent (Figure 3.12), a student of Bartlett, presented a model that formalised and represented pictorially the putative processing stages in perception, memory and learning, and it was addressed to the realistic activity of communication. Thus, Broadbent combined Bartlett's approach of examining skilled tasks with Craik's modelling metaphor.

When Bartlett and Craik proposed their theories of perception relatively little was known about the brain mechanisms that mediate perception. This is one of the reasons why the Gestalt psychologists were able to propose their speculative neurophysiology of vision. Craik anticipated that the computer would be a powerful tool to simulate theories of perception, as well as providing a metaphor for the processes of perception and cognition themselves. Since the late 1960s, the study of visual perception had been profoundly influenced by computers. As well as allowing scientists to collect or to analyse data more quickly, the digital computer provided a tool for the laboratory scientist to develop new ways of testing the visual system with novel kinds of visual displays.

Richard Gregory (Figure 3.12) also studied under Bartlett and perception, especially when it was illusory, was the area in which he specialised. He was fascinated by perceptual paradoxes and sought ways of casting light on them. He categorised and interpreted visual illusions; the categories were generally optical, physiological and cognitive and interpretations of the last were in terms of the Helmholtzian concept of unconscious inference or, as he put it, perceptual hypotheses. His theoretical position was grounded in British empiricist philosophy, and Helmholtz was seen as its ablest exponent in the context of vision. Gregory proposed that illusions were seen because inappropriate hypotheses were applied to the patterns of stimulation so that two-dimensional representations were interpreted with respect to their three-dimensional referents. The influence of cognitive processes on perceptual experiences defined his approach and he applied it to a wide range of phenomena. He wrote: "Perceptions are considered as being *predictive hypotheses* – essentially similar to hypotheses in science".

Figure 3.12 Left, **Schema** Frederic Charles Bartlett (1886–1969) emphasised the constructive aspects of memory and perception as well as introducing the concept of 'schemas' into psychology. The text in which his portrait is embedded is taken from his book *Remembering*, and it describes the concept of 'schema'. Centre, **Channel capacity** shows Donald Eric Broadbent (1926–1993) who was one of Bartlett's students; he analysed dichotic listening and linked perception more closely with communication. He is shown in the context of a task he examined in great detail: several messages are presented simultaneously, but here in vision rather than hearing. The channel capacity of the observer is exceeded by these signals so that processing has to be limited to one at once – if you can see the face then the text cannot be read. Right, **Tile illusion** Richard Gregory (1923–2010) examined visual illusions, like the distorted squares that carry his portrait, and applied cognitive interpretations to illusions, and to vision in general

A similar cognitive revolution took place in America, but a little later. The approach to perception adopted by Bartlett was applied to human operators of complex systems. The experimental research on perception in the 1940s harmonised with developments in cybernetics. Information theory was developed in the context of telecommunications, and the mathematical measurement of information was formalised in the late 1940s; its powerful impact on perception was felt in the 1950s. George Miller (Figure 3.13) linked the concept of limited information capacity to absolute perceptual judgements. He also allied the processes of perception more closely to those of language, and this was amplified by Noam Chomsky (Figure 3.13), a linguist who has introduced a wide range of novel terms into the analysis of language. In Chomsky's transformational grammar, one of the principal distinctions is between the surface and deep structures of a sentence: the surface structure corresponds to the sequence of words as written or spoken, whereas the deep structure refers to their underlying meaning.

Fred Attneave (1919–1991) devised procedures to determine the locations of highest information in simple patterns. They corresponded to boundaries of brightness (contours) and particularly to abrupt changes in contour direction (corners). Support for the significance of contours in perception derived from two other sources – single unit recordings from various levels in the visual pathway, and scanning eye movements. Indeed, early attempts to stabilise the retinal image by compensating for any involuntary eye movements resulted in disappearance of the target. Theoretical attention shifted towards pattern recognition by both humans and computers because they were both thought of as information

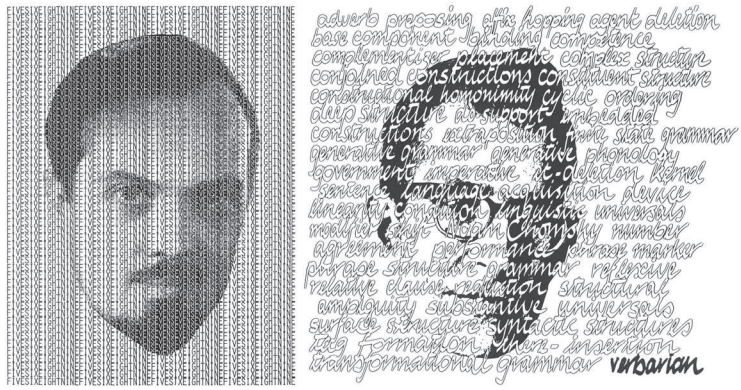


Figure 3.13. The languages of perception. Left, **The magical number seven, plus or minus two** George Armitage Miller is defined by his magical number and surrounded by its neighbours. Right, **Deep structure** Noam Chomsky is enclosed within terms he coined; the surface structure of the list is not punctuated, but the terms can be suitably segmented with respect to its deep structure: although the individual words are not alphabetically arranged, the novel terms are.

processors or manipulators of symbolic information. The patterns were typically outline figures or alphanumeric symbols and rival theories, based on template matching and feature analysis, vied for simulated supremacy at recognition and one result was a model named pandemonium! Stuart Sutherland (1927–1998) sounded a cautionary note on this endeavour that was not generally heeded then nor has been subsequently: "Patterns are of importance to animals and man only in so far as they signify objects. It is the recognition of objects that is vital for survival and as a guide to action, and the patterned stimulation of our receptors is of use only because it is possible to construct from it the nature of the object from which it emanated".

Despite the use of the machine metaphor, Craik did formalise the input, processing, and output components of servo-systems in a manner that could be applied to digital computers. When computing machines increased in speed and complexity, the tasks that they could simulate became more complex. Concepts from engineering, like information and self-organisation, were also integrated with a growing knowledge of neurophysiology, resulting in the computer becoming a metaphor for the brain.

Bartlett's emphasis on the constructive nature of perception found an echo in America in the 'New Look' experiments, where motivation was considered to interact with perception. Similar experimental investigations had been undertaken earlier by Brunswik who examined the perceived sizes of postage stamps of different value. Ames' many demonstrations of the ambiguities of stimulation and their perceptual resolutions were also accorded renewed attention in this cognitive climate. In those heady postwar years personality flirted with perception, but their liaison was not lasting. Certain subthreshold recognition phenomena were brought to their perceptual defence, but the sober verdict was not in their favour:

"It would seem wise to regard with great caution the existence of limitations on speed and accuracy of perception imposed by personality factors, at least in normal observers".

Craik, as well as Turing, anticipated that the computer would be a powerful tool to simulate theories of perception, as well as providing a metaphor for the processes of perception and cognition themselves. Since the late 1960s, the study of visual perception had been profoundly influenced by computers. As well as allowing scientists to collect or to analyse data more quickly, the digital computer provided a tool for the laboratory scientist to develop new ways of testing the visual system with novel kinds of visual displays. The move away from reliance on oscilloscopes to present sine wave and other simple patterns facilitated the increasing use of more naturalistic patterns, as well as those which can be constructed and manipulated in controlled ways. Computer developments also enabled the better recording of eye movements and the linkage of eye movements to changes in display features, allowing a number of groups to conduct ingenious experiments into the control of eye movements in reading. Computer developments also created new problems. Unknown non-linearities in screen display properties or their temporal characteristics presented problems that were not initially foreseen, in addition to a variety of problems that have been termed visual stress.

During the 1970s, David Marr (see Figure 1.8) set out to develop a complete framework for vision, spanning the very lowest level processes within the retina up to the process of visual object recognition. The key feature of Marr's theory was that vision can be understood at different levels. The first 'computational' level is a theory of the task that the visual system is to solve, and an understanding of the constraints that can enable solution of that task. The second level, of 'representation and algorithm', is a means of achieving the task, and the final 'hardware implementation' level describes how the brain, or a computer, actually implements these algorithms in neural tissue or silicon.

In addition to presenting a unified approach to different topics within vision, Marr and his colleagues also presented a theory of the different stages of representation involved in the interpretation of a retinal image. In so doing, Marr distinguished a stage which made explicit the three-dimensional layout of the world with respect to the viewer (the 2½-D sketch), potentially useful for action in the world from the more abstract 3-D models which allowed object recognition. Despite this distinction, however, his framework fails to anticipate the more recent idea of a separation of visual pathways for action and for object categorisation. Marr's framework for visual object recognition was developed and extended by many others. Subsequent theories about object recognition, however, departed from Marr's economical object-centred representations, to develop representational ideas based around the representation of multiple viewpoints. Thus, it is not the details of Marr's theory which have so far stood the test of time, but the approach itself.

Another key feature of Marr's theory was the principle of modularity: "Computer scientists call the separate pieces of a process its modules, and the

idea that a large computation can be split up and implemented as a collection of parts that are as nearly independent of one another as the overall task allows, is so important that I was moved to elevate it to a principle: the principle of modular design".

Indirect and direct perception

All theories of perception considered so far are what can be called indirect; some mediation between the pattern of stimulation, its effects on the visual system and perceptual experience was implicit – some mode of representation. The term often applied to this was that perception went beyond the information given. This is at the heart of empiricist approaches to perception and is in the organisational principles of Gestalt psychology. An alternative approach was advocated by Gibson (see Figure 1.1). Rather than considering the senses as independent of one another he integrated them into perceptual systems in which time was an essential component. Gibson retained separate perceptual systems which he called orienting, auditory, haptic-somatic, tasting and smelling, and visual. His ideas established a new field of 'ecological' optics which has been tilled by many in recent years.

Gibson sought to stem the cognitive current and developed a novel theory that owed more to the philosopher Thomas Reid (see Figure 3.3) than to his own contemporaries: "When the senses are considered as perceptual systems, all theories of perception become at one stroke unnecessary. It is no longer a question of how the mind operates on the deliverances of sense, or how past experience can organize the data, or even how the brain can process the inputs of the nerves, but simply how information is picked up". Gibson abolished the senses when he replaced them by perceptual systems. That is, the distinction between sensation and perception was abandoned, and perceptual systems afforded useful information for interaction with the external world: "We shall have to conceive the external senses in a new way, as active rather than passive, as systems rather than channels, and as interrelated rather than mutually exclusive. If they function to pick up information, not simply to arouse sensations, this function should be denoted by a different term. They will here be called *perceptual systems*". Moreover, there was considered to be a perfect correlation between the stimulus and its perception; no stages of representation were involved in perception. The doctrine of specific nerve energies, that had informed almost all studies of the senses since the time of Johannes Müller, emphasised the indirectness of perception. The brain had access only to the nerve signals initiated by external objects, not to the objects themselves. Gibson cast aside this tradition in favour of direct perception. Despite his pejorative purview of conventional perceptual experiments, the strongest support for his position derived therefrom: simplified dynamic dot patterns could be recognised far more easily than static ones.

Gibson called for neither physiological nor computational support for his theory, but one of his contemporaries, Donald Hebb (Figure 3.14), was integrating both of these. Hebb, in his speculative synthesis of perception and learning,

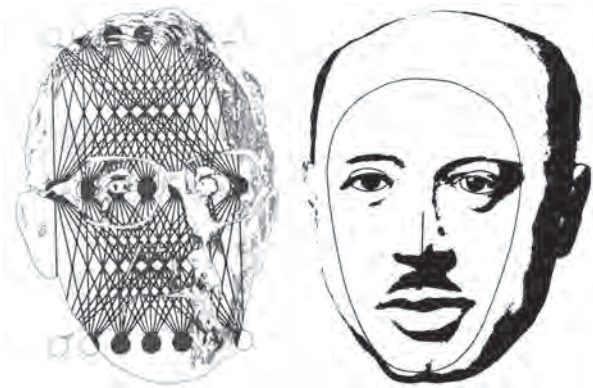


Figure 3.14 Left, **Hidden units** Donald Olding Hebb (1904–1985) can be seen in the middle (hidden) layer of a three-level network: each cell in the middle layer is connected to every other one in the levels above and below it. Right, **Schematic face** Egon Brunswik (1903–1955) conducted studies on the perception of schematic faces and is portrayed with one that reflects his own soft, sad, and intellectual features.

wove patterns with networks of neurons connecting perception to its underlying physiology. He proposed that perceptual learning takes place when assemblies of cells fire together; their reverberating activity resulted in synaptic changes which further increased the probability of the nerves firing together. The functions of cell-assemblies and phase sequences were based on his neurophysiological postulate: "When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased". Hebb later applied the concepts to account for a wide range of phenomena, from stabilised retinal images to sensory deprivation.

Another indirect theory was proposed by Egon Brunswik (Figure 3.14) who introduced probabilistic functionalism into psychology. His was a functionalist theory because it emphasised the adaptive nature of behaviour with respect to objects in the environment; it was probabilistic because behaviour in an unpredictable environment must be based on the statistical regularities that occur within it. He was influenced by the Vienna circle of logical positivist philosophers and believed that the probabilistic methods applied to the physical sciences were appropriate to psychology, too. He applied probabilistic functionalism principally in the area of perception. He was concerned with how we derive veridical information about objects; how the distal stimulus is perceived as constant despite wide variations in the proximal stimulus. Accordingly, he examined perceptual constancy and devised a formula for assessing it, now known as the Brunswik ratio. Veridical perception was based on the use of a family of cues that differ in their ecological validity – the correlation between proximal cues and distal stimulus. By experience, greater statistical weight was placed on the cues with high ecological validity.

The physiological dimension

Most historical investigations of eyes have been addressed to the image-forming variety of vertebrates. This was so despite the knowledge of many other evolutionary adaptations for transducing light energy. Until the neuron doctrine was firmly established in the 1890s, examination of invertebrate eyes was not considered to offer insights into vertebrate visual processing. The situation was transformed by Haldan Keffer Hartline (1903–1983) with his studies of responses to light in the horseshoe crab, *Limulus polyphemus*. His pioneering neurophysiological experiments, together with allied investigations of vertebrate visual responses, resulted in the emergence of a new conception of retinal processing: lateral and recurrent interactions, occurring in a complex network of neuronal circuits, were considered fundamental mechanisms for visual information processing.

In the last few decades there have been major strides in furthering our understanding of neural processes in the visual system. These discoveries have been taken to support the view that vision involves a sequence of stages in which different aspects of the stimulus, like colour, contour and motion, are extracted. Research on patterned stimulation at the receptor level had proceeded throughout the first half of the twentieth century, but its pace quickened thereafter. The glimmerings of pattern processing beyond the receptors emerged in the 1950s, and were amplified in the 1960s. When recordings of nerve impulses could be made from individual cells in the visual pathway, their adequate stimuli could be determined (see Chapter 5). It came as something of a surprise that retinal ganglion cells of frog responded to quite complex features of stimulation (like moving dark regions of a specific visual angle, resembling a bug), and stimulus properties that excited or inhibited neurons were generally called 'trigger features'. Retinal ganglion cells of cat, on the other hand, were excited by rather simpler stimulus arrangements. It was found that they were concentrically and antagonistically organised; if the centre was excited by light, the surround was inhibited, and vice versa. Such an arrangement served the detection of differences in luminance well, but steady states would have little effect, since excitation nullified inhibition. This pattern of neural activity was retained in the lateral geniculate body, but it underwent a radical change at the level of the visual cortex. From the 1960s, David Hunter Hubel and Torsten Nils Wiesel (who are both shown in Figure 3.15) found that single cells in primary visual cortex (V1), first of cat then of monkey, responded to specifically oriented edges.

Physiologists refined the stimulus characteristics of trigger features while psychologists sought their phenomenal counterparts. Almost any experiment involving contours paid lip service to Hubel and Wiesel, despite the tenuousness of the links between particular phenomena and their underlying physiology. At least an appeal to trigger features was considered preferable to reliance on the speculative neurophysiology advanced by Gestalt psychologists. The concept of channels or spatial filters emerged during the 1960s, and it was applied with particular rigour by Fergus Campbell (1924–1993) and his colleagues to the detection of and adaptation to sine-wave gratings. The attraction of gratings was

Figure 3.15 **Feature detectives** David Hunter Hubel and Torsten Nils Wiesel are shown enclosed within the pattern of ocular dominance columns that they disclosed in monkey cortex using autoradiography. Hubel and Wiesel were awarded the Nobel Prize in Physiology or Medicine in 1981.

that they provided at one and the same time a definition of the stimulus and theory of the response to it.

The 1960s saw the beginnings of a split between a 'cognitive' approach, where the goal of vision could be seen as an abstract categorisation of the objects of vision, and an 'action' approach, where vision was part of an integrated system allowing manipulation of and navigation through the world. This distinction has matured in contemporary approaches to vision, both through the influence of Marr, and through further developments in neuroscience and neuropsychology.

The growth in visual neuroscience has revealed an increasingly complex, though elegant, picture, and diagrams of visual areas, their interconnections, and their microstructure are likely to get increasingly complicated. Evidence for the analytic separation of different aspects of the visual scene – motions, forms, colours, etc., raises the question of how these elements become associated, or 'bound'. In addition to the problem of binding within the visual domain, however, there is the problem of how and when different modalities (vision, touch, hearing) become integrated or otherwise influence each other.

Illusions have been used to provide signs for the neuroscientists to pursue. It is doubtful whether neuroscientists can provide signs to direct research on illusions until there is an adequate neurophysiological theory of visual processing. At present, there are sets of sub-theories which are over-interpreted in terms of visual psychophysics. It might seem unreasonable to demand a theory of neuroscience before applying it to visual perception, but it is realistic. It will take a long time, but that might be shorter than making many false starts, as happened in the last few decades of the twentieth century. It would also have the positive effect of making us concentrate on perception rather than its putative underpinnings. The conundrum is that there is a demand to conduct theory-driven experiments, and most of the theories are based on inadequate neuroscience. The need for a good perceptual taxonomy becomes pressing under such circumstances.

Twenty-first-century prospects

The origins of psychology are to be found in philosophy and medicine, and these two disciplines complement one another in the study of perception. Some of the dominant themes in philosophy, like rationalism and empiricism, have drawn support from perception and have addressed some of the enduring problems that are still under study. The origins of psychology as an independent discipline were rooted in perception. Wundt enlisted the psychophysical methods developed by his colleagues at Leipzig, Weber and Fechner; he also adopted the instruments for the laboratory study of perception, like chronoscopes, stroboscopes, stereoscopes and tachistoscopes. While these techniques continued to be used throughout the twentieth century, Wundt's method for studying consciousness, analytic introspection, did not survive the test of time. Opposition from behaviourists, on the one hand, and Gestalt psychologists, on the other, resulted in its demise; only the latter was grounded in perception. Sensory processes were examined by physiologists throughout these psychological skirmishes; with the emergence of microelectrode recordings from single cells (in the second half of the twentieth century) physiological theories became more prominent. Their emergence coincided with the applications of information theory to perception and cognition. Indeed, cognitive theories of perception were eagerly embraced, particularly when they could be described in computational terms. The computational approaches have insinuated themselves in the language used to characterise processes in perception. Computer terminology has been adopted to distinguish between 'bottom up' and 'top down' theories. The former try to interpret phenomena in terms of the early stages of visual processing and they place little emphasis on cognitive involvement in perception. The opposite 'top down' theories suggest that our expectations and cognitive strategies determine what we perceive. They refer back to Helmholtzian speculations about unconscious inference.

Perceptual theories in the twenty-first century are developing along many lines. Perhaps the two most prominent are the trends towards mathematics and mental modelling. The first is expressed in several strands. On the one hand, emphasis on inferential processes in perception, exemplified by Brunswik's probabilistic functionalism, finds more concrete expression in Bayesian approaches to perception. These have found a new champion in an old mathematician. Rev. Thomas Bayes (Figure 3.16) provided a formulation of inference that is now widely used in perceptual theories. According to Bayesian theorists, there is always some prior probability of encountering certain events and the weighting of these is modified by perception. The virtue of this approach is that it is quantitative, as long as values can be applied to the prior probabilities, and the theory has wide application in computer vision. On the other hand, neo-Gibsonians apply increasingly sophisticated analyses of stimuli that change over time (events), particularly in naturalistic settings. Mental modelling is reflected in the imaging and computational strands. Advances in brain imaging, and the computer software associated with it, have resulted in proposals for localising perceptual functions at more specific brain sites. Computational models of both pattern recognition and

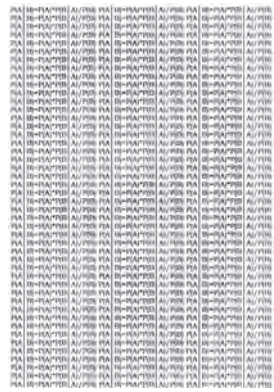


Figure 3.16 **Bayesian inference** Possibly a portrait of Thomas Bayes (c. 1701–1761). There is no known authenticated portrait of Bayes, but the one used in producing this illustration is most commonly presented as a likeness of him. The equation $P(A \mid B) = P(A) * P(B \mid A)/P(B)$ is Bayes' theorem. In the context of perceptual theory, the terms are (A) hypothesis and (B) evidence regarding its validity. Thus, the probability that the hypothesis will be supported by the evidence is determined by the prior probability P(A) and the ensuing perception. If the hypothesis is supported the prior is strengthened and vice versa.

robotic control inform as well as being driven by perceptual theory. Perception, particularly visual perception, will increasingly be integrated with motor activity and this will be reflected in both theory and experiment.

All theories of perception mentioned in this chapter continue to be pursued, but the terminology in which they are shrouded has changed. Almost all areas of the life sciences are now prefixed by neuro- and they tend to exist under the umbrella of 'neuroscience'. This has applied particularly to imaging brain events. Neuroimaging has developed apace and novel techniques are being devised. This has led to increased attempts to localise visual functions in specific brain regions.

The grand theories of the twentieth century have given way to more specific variations of them in the present century. One trend is for greater technical application to the problems of vision, as in neuroimaging and eye movement research. There is also a move towards increased quantification of the methods and models that are applied to phenomena. Despite these trends, certain fundamental aspects of vision remain and these need to be addressed by all theories like issues concerning frames of reference which will be the topics of Chapters 6 to 8. Before they are discussed, the optical and physiological dimensions of vision will be examined in more detail.

Reference notes

Aristotle's descriptions of afterimages and aftereffects are in volume 3 of *The works of Aristotle* (Ross, 1931). Boyle's account of the ill-advised observation of the sun through a telescope is recounted in his *Works* (Birch, 1966). There are many histories of medicine which describe the development of light and sight

from ancient times. Polyak (1957) contains a detailed history of advances in gross and microscopic anatomy of the eye, as does Wade (1998a). The development of ideas about the nature of light and optics is covered by Darrigol (2012), Park (1997) and Ronchi (1970). Finger (1994) places the growth of understanding of the visual process in the broader context of cognitive development and its relation to the brain. A more selective, but very readable, account of the growth of ideas about vision and the brain can be found in the essays by Gross (1998). The motion aftereffect has been examined since the time of Aristotle, and its history and current status can be read in Mather, Verstraten and Anstis (1998). A detailed history of psychological theories in all areas, including perception, is provided in Rieber (2012); entries for the individuals mentioned in this chapter can also be found there.

There are many books on the history of philosophy, but its impact on the history of psychology is cogently covered in Robinson (2008). The quotation from Locke on empiricism is from his *Essay concerning human understanding*. An English translation of part of Goethe's *Theory of colours* is available (Goethe, 1970), as is one for Purkinje's book on subjective visual phenomena (Wade and Brožek, 2001). Helmholtz is a towering figure in the history of vision and his wider contributions to the dominance of German science in the nineteenth century are described in Cahan (1993). The convoluted controversy between Helmholtz and Hering over empiricism and nativism is traced in sharp detail by Turner (1994) and also by Howard (1999). The quotation from Helmholtz regarding 'tokens of sensation' is taken from the translation of his *Handbuch* into English (Helmholtz, 1925/2000). Hering's quotation regarding binocular visual direction is from Hering (1942). The authors have examined a neglected aspect of Helmholtz's sensori-motor coordination with respect to golf (Wade and Swanston, 2001).

Wundt (1874) called his New Psychology 'physiological psychology' and the history of the institutionalisation of the subject can be found in Danziger (1990) and in Rieber and Robinson (2001). Titchener both translated Wundt into English as well as adding his own interpretations of perception, to which he gave the name 'structuralism' (Titchener, 1910). The behaviourist manifesto quotation is from Watson (1913). Gestalt psychologists wrote extensively (see Köhler, 1930, 1940) and they were extensively written about. Translated readings from many of the early papers are available in Ellis (1938). The quotation about the difficulty of recognising pictures of faces when they are inverted was in Köhler (1940, pp. 25–26). Ash (1995) places the Gestalt movement in a broader historical context and provides good coverage of their early perceptual experiments. Artistic precursors of Gestalt grouping principles can be seen in Roman mosaics, and these are described in Wade (2012b).

The cognitive approach has its seeds in Bartlett's (1932) work on memory and skilled activity (see Saito, 2000); the movement spread more widely in the 1950s. Bartlett's students Craik (1943) and Broadbent (1958) developed the machine metaphor and the links between perception and communication generally. The quotation regarding the filter model is from Broadbent (1958, pp. 306–307). Sutherland's quotation regarding patterns is from Sutherland (1973, p. 157). The

cautionary note about linking personality and perception was voiced by Vernon (1970, p. 237). Seeing, and the instruments that can be used to extend its ambit, were Richard Gregory's constant concerns, and his ability to communicate his fascination inspired many students to pursue the path of perception. His historical approach to the vision was expressed in *Eye and brain*, first published in 1966 and still thriving in its fifth edition, having been translated into twelve languages. His final book, *Seeing through illusions* (2009), displayed his perception and his delight in word play. Miller, Gallanter and Pribram (1960) captured the spirit of the early cognitive times in America. The 'cognitive revolution' has been charted by Gardner (1987) and a personal view of one who was at its origins is given by Miller (2003). Marr's computational approach to vision was published after his untimely death (Marr, 1982). Its importance to what is now called computational neuroscience is evident by the fact that it has recently been republished (Marr, 2010). There is now a Marr Prize that is awarded annually for outstanding contributions to computer vision.

Gibson wrote three books on perception (Gibson, 1950, 1966, 1979) in which his evolving analysis of direct perception was presented. The quotations regarding information pickup and perceptual systems are from the 1966 book (pp. 319 and 47, respectively). Hebb's first book appeared at about the same time as Gibson's, and it was in it that he gave his neurophysiological postulate (Hebb, 1949, p. 62). Brunswik had been ploughing the perceptual field for much longer (Brunswik, 1934); the schematic faces were illustrated in Brunswik (1956). Hebb did frame his theory in physiological terminology, but the great advances in single unit recordings were made after his book was published. Hubel and Wiesel (2005) have reprinted their papers on single unit recordings from the visual pathways.

There is a multitude of books on neuroscience, and many of them make contact with the senses. The physiological underpinnings of vision are expounded by Fahle and Greenlee (2003) and Daw (2012). Kanwisher and Duncan (2004) edited a book on neuroimaging and visual cognition, and Smith *et al.* (2008) edited a special issue of the *Journal of Vision* devoted to neuroimaging in vision science. A historical account of Bayes' theorem is provided by Cowles (1989) and the theoretical links between Helmholtz and Bayes are explored by Westheimer (2008). Gregory (2006) examined the links between Bayes' theory and visual illusions. Contemporary accounts of vision embracing Bayesian theory are given in Purves and Lotto (2010).

4 Optics and the eye

Over four centuries ago, Porta (see Figure 2.21) considered he had solved the problem of vision! In a book entitled *Natural magick* he wrote: “Hence it may appear to Philosophers, and those that study Opticks, how vision is made; and the question of intromission is taken away, that was antiently so discussed; nor can there be any better way to demonstrate both, than this. The Image is let in by the pupil, as by the hole in the window; and that part of the Sphere, that is set in the middle of the eye, stands in stead of a Crystal Table.” Porta described how a camera obscura (a dark chamber), with a lens in the enlarged aperture, could project an image of an external scene onto a suitably located screen (Figure 4.1). He was not the first to describe the camera obscura: Ibn al-Haytham (Alhazen) had observed an inverted image in a dark chamber, and pinhole experiments were conducted by late medieval students of optics. Nor was Porta the first to equate its function with the eye; Leonardo da Vinci conducted experiments with a camera

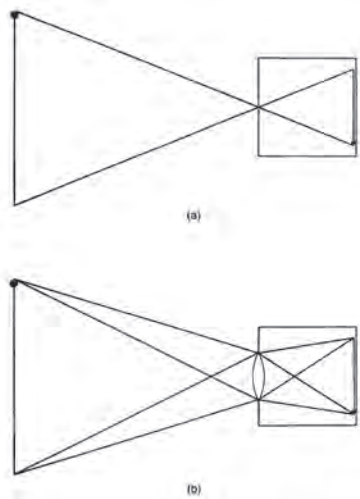


Figure 4.1 (a) The formation of an image in a camera obscura. The pinhole aperture allows certain light rays from objects in the environment to enter the camera, where they form an inverted image on the screen. In (b) the addition of a lens increases the proportion of the rays from a point on the object which are brought to a single point in the camera.

Figure 4.2 Revolutionaries of visual optics: Left, **Optical astronomer** Johannes Kepler (1571–1630) and the title page of his book on optics in which he describes how images are formed in the eye. Right, **Ocular astronomer** Christoph Scheiner (1573–1650) and the frontispiece of his book on the eye in which its correct anatomy was represented.

obscura, and drew an analogy between its operation and that of the eye. Porta's quotation indicates the power of the analogy between eye and camera. If the eye operates optically like a camera then the processes in both are likely to be similar. The eye–camera analogy still influences many analyses of vision.

Porta was able to use the demonstration to argue against extramission theories of light and sight – that light was produced within the eye and projected outwards. Image and eye were fused by Porta, but the retina did not enter into his account. As is evident from the quotation, he echoed the ancient view that the lens was the 'seat of vision' and hence the screen onto which images were projected.

Porta stood at the threshold of a revolution in visual optics, but he was not able to cross it. He lacked an adequate understanding of image formation in the eye as well as an accurate knowledge of the anatomy of the eye. Kepler (Figure 4.2) was to provide the first of these in 1604 and Christoph Scheiner (Figure 4.2) supplied the second in 1619. Moreover, the nature of light itself was a mystery to Porta. As noted in Chapter 2, in his day, there were three categories of light that were deemed worthy of study: optics or direct vision, catoptrics or reflections and dioptrics or refractions.

Light is the stimulus for vision and the eye is the organ which responds to light energy. The environment contains objects which can emit, reflect or absorb electromagnetic radiation. Such radiation is ubiquitous, but only a small portion can be directly sensed by living organisms, and a smaller part still can be detected by the eye. The nature of light is now thought of as a phenomenon of physics, but those who advanced its science also added to our knowledge of sight. In the late seventeenth century Newton wrote: "I understand light to be any entity which proceeds directly from a bright body and is able to excite vision". The entities to which he was alluding were corpuscles or particles of light that bounced through the ether. Those present in white light could be decomposed into the spectral components by refraction through a prism, which in turn could

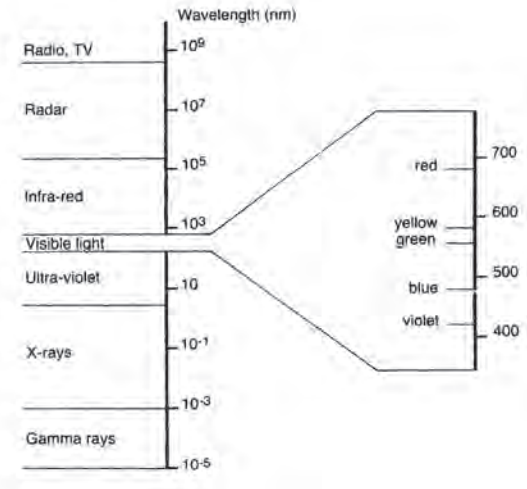


Figure 4.3 Part of the electromagnetic spectrum, showing the small proportion which corresponds to visible light. The expanded section indicates the approximate wavelengths at which certain colours are normally seen.

be recombined to produce white light. Newton questioned the purity of white but he also appreciated the subjectivity of colour perception: the rays, he said, are not coloured. However, the phenomena of diffraction and interference could be much more readily interpreted using the concept of waves. Furthermore, as a consequence of Maxwell's theoretical work in the nineteenth century, light is now described as electromagnetic radiation – a wave having both magnetic and electrical fields – but so are cosmic rays and radio waves. Light is distinguished from other electromagnetic radiations on the basis of sight: that small part of the spectrum to which the human eye is sensitive is called light.

At the beginning of the twentieth century the principal question was: Is light a particle or is it an oscillation? Evidence could be adduced to support both, and there remained an uneasy ambiguity between Newtonian corpuscles – now called photons – and Maxwell's electromagnetic waves. They have been unified mathematically, in quantum field theory, but this has not transferred readily to experiments on light, or our conceptions of it. Depending on the aspects of light that are measured it can display the properties of electromagnetic waves or photons. The light we experience is typically reflected from objects in the environment – light renders those objects accessible to visual experience, and hence enhances our chances of survival. Classical physics (up to the time of Young) examined the properties of light using visible phenomena. Particles and waves are conceivable properties of light as they correspond with perceivable aspects of our experience: larger objects can be divided into smaller and smaller parts and waves are readily seen in patterns of disturbances of water. Modern mathematical analyses of light render its characteristics inconceivable as well as imperceptible.

Electromagnetic radiation can be considered as a wave, and as such can vary in wavelength. Wavelength is defined by the distance between successive peaks in the wave, and for light it is measured in nanometres (nm), where 1 nm equals one billionth of a metre (10^{-9} m). The range of human vision extends from around 380 nm to around 720 nm, and this band of electromagnetic radiation is referred to as light. To someone with normal colour vision, light of different wavelengths appears coloured making up the visible spectrum (Figure 4.3).

Light sources in the natural world are relatively rare, and are limited to celestial bodies such as the sun and stars, lightning, fire, and biochemical processes in living organisms (bioluminescence). Of these the sun is by far the most important, and its location provides a constraint on the appearance of illuminated objects that plays a fundamental role in the perception of depth and shape. The interpretation of three-dimensional structure from shadows seems to be based on an assumption that light comes from the sky above. Our widespread use of artificial sources of illumination may cause errors in the perception of objects, if the light source is located below eye level. Objects which are not light sources can only reflect some part of the light which falls upon them. Generally, objects are neither perfect reflectors nor perfect absorbers of light. Their molecular structure causes certain wavelengths in the incident radiation to be absorbed, and others to be reflected. For example, a blue object appears to be this colour because it absorbs most of the incident light energy with wavelengths above around 550 nm. In practice, the perceived colour of objects is also influenced by factors such as the colour of adjacent objects and of the incident light. The light reflected from one object may in turn fall upon other objects, with their own reflective properties. Our environment is therefore filled with emitted and reflected light, which forms a field of energy carrying information about the environment's characteristics. In order to see, we must first capture a sample of this radiation, in such a way that the information it carries is not too distorted or degraded.

Image formation in the eye

The visual revolution commenced four centuries ago, shortly after Porta wrote the words at the head of this chapter. Kepler described the manner in which images are formed in the eye before its detailed anatomy had been accurately described and illustrated. The image on the retina was inverted and reversed, which was seen as posing a problem for vision, and one that Kepler was not prepared to speculate upon. This problem puzzled many students of vision, even after its solution had been proposed by Molyneux (see Chapter 2). He argued that the terms erect and inverted are only relative to the direction of gravity; moreover, the process of perception involves the mind as well as the eyes. It took most of the seventeenth century for the distinction between the retinal image and perception to be appreciated.

Following Kepler's insight, students of vision tended to reduce the analysis of vision to an analysis of the image formed in the eye. That is, vision became a problem for geometrical optics, and there was a concrete model for vision (the

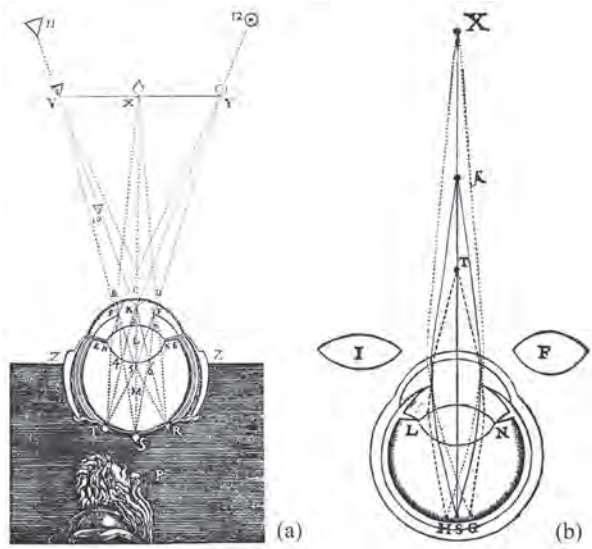


Figure 4.4 Descartes' representations of (a) image formation in the eye and (b) his speculations regarding changes in lens curvature and accommodation (from Descartes' *Dioptrique*).

camera obscura with a lens) which could be used in an experimental way. In 1619, Scheiner was the first to provide an accurate diagram of the anatomy of the eye, and he carried out experiments with an excised eye of an ox: when the rear coats were cut away (leaving the dark retina intact) and the eye was placed in a small hole of a dark room an inverted image of the scene outside could be observed. The most famous depiction of this was by René Descartes in his *Dioptrique*. It is shown in Figure 4.4a: rays of light from the object VXY are refracted by the cornea (BCD) and crystalline lens (L) to focus an image RST on the retina.

Accommodation

The analogy between eye and camera introduced a new set of problems in the study of vision. If the camera can only focus on objects at a particular distance, how is the eye able to focus upon objects over a wide range of distances? This is the problem of accommodation. From the time of Kepler to the middle of the nineteenth century accommodation was one of the most intensively studied and controversial topics in vision. Not surprisingly, since the equation of the eye with a camera had proved so popular, the solutions were often derived from characteristics of cameras. A camera with a small aperture has a much greater depth of focus than one with a larger aperture; moving the camera lens towards or away from the screen onto which images are projected will vary the distance at which objects are sharply focused; conversely, moving the screen itself will have

the same effect. Each of these physical speculations was advanced, together with others that were physiological.

Despite the analogies of accommodation with focusing in a camera, Descartes' physiological speculations were to prove particularly astute (see Figure 4.4). The lens itself was considered to change its curvature, becoming more convex for focusing on near objects, and less convex for more distant ones. He even suggested that accommodation provides a source of distance information for objects that are close to the eye. The strongest support for the involvement of the lens in accommodation came from an eighteenth-century Scottish physician, William Porterfield (1696–1771); he examined a patient who had had the lens removed from one eye (because of cataract) and was unable to accommodate at all without the aid of a convex lens; moreover, the power of the lens required to be modified for objects at different distances.

Thus, writers on the eye and vision selected one or more of these hypotheses as their candidates for accommodation until the late eighteenth century, when Young reported his experimental enquiries supporting changes in lens curvature. Such support was not derived from direct evidence, but rather from the rejection of all alternative hypotheses. Changes in corneal curvature were excluded in two ways: the images of candle flames reflected from the cornea did not change with variations in accommodation, and immersion of the eye in water did not abolish accommodation. Elongation of the eye was rendered untenable because accommodation was still possible when considerable external pressure was applied to the eye. The mechanism by which changes in lens curvature could be achieved was proposed by Helmholtz. In order to measure the curvatures of the optical surfaces in the living eye Helmholtz invented the ophthalmometer; he confirmed the speculations of Descartes and the experiments of Young that the lens changes curvature during accommodation, and he proposed the mechanism by which this is achieved. Helmholtz wrote: "On contraction, the ciliary muscle could pull the posterior end of the zonule forwards nearer the lens and reduce the tension of the zonule. If the pull of the zonule is relaxed in accommodating for near vision, the equatorial diameter of the lens will diminish, and the lens will get thicker in the middle, both surfaces becoming more curved".

The accommodation of the eye is driven by a reflex response to the retinal image, but the basis of this is not well understood. It does not seem to depend simply on the degree of blur, and is influenced by other factors such as the wavelength of the light. Feedback from the state of accommodation can act as a source of information for the visual system about the distance of objects; the more the accommodation, the closer the object. Clearly, this is limited to the range of distances between the far and near points, and will not convey much information in good illumination, when less accommodative effort is required. In experiments on depth perception it is a common practice to have observers look through an artificial pupil, which is a small aperture about 1 mm in diameter placed immediately in front of the eye. This has the effect of ensuring that the retinal image is in focus regardless of the state of accommodation, which is held constant as a factor in determining the perception of distance.

Ocular anatomy

Some knowledge of the anatomy of the eye must have been available several thousand years ago because there were not only medical practitioners in Babylonia, Mesopotamia, and Egypt but also eye specialists; they must have had a working knowledge of ocular anatomy in order to carry out the operations they are known to have performed. Their skills and understanding would have been passed on to Greek physicians, who both recorded and developed them. Many Greek texts, through their translations, have been transmitted to us, but any illustrations that they might have included have not survived.

Printed figures of the eye were published from the beginning of the sixteenth century, but they were derived from dogma rather than dissection. Following the relaxation of sanctions prohibiting dissection of human bodies in the fourteenth century, knowledge concerning anatomy in general slowly began to be based on more secure ground. The dissecting skills of the anatomist were critical, and the major advances came with practitioners like Leonardo and Vesalius. Leonardo's detailed drawings of dissections did not make any immediate impact because they remained both in manuscript form and in private hands. Unlike his anatomical drawings of the musculature, those of the eye reflected a conflation of dissection and dogma: his rather crude drawings reflected a reliance on Galen, even though he did prepare the excised eye (by boiling it in the white of an egg) for dissection. His drawings of the eye showed the lens as spherical and central in the eye, and the optic nerves passed to the cerebral ventricles.

The renaissance of anatomy is associated with Vesalius, who published his book *De humani corporis babrica* (On the structure of the human body) in 1543. It is taken to be a synthesis of science and art because of the high quality of the anatomical illustrations. His diagram of the eye did not match the detail or accuracy of those for the skeletal musculature and internal organs: a symmetrical lens was still located in the centre of the eye and the optic nerve was situated on the optic axis. He listed the various structures, but did not pursue their function in any detail. Throughout the sixteenth century diagrams of the eye showed increasing refinement, with more precise locations of the various structures within it. In 1583, the anatomist Felix Platter (1536–1614) stated that the retina rather than the lens was the receptive organ for vision, which received light through the lens, and which could discriminate colour and form; his diagram of the eye (Figure 4.5a) was reproduced in Kepler's book on optics. The first accurate diagram of the eye (Figure 4.5b) was that by Scheiner; the lens and its curvatures are appropriately represented and the optic nerve leaves the eye nasally. Scheiner's analysis was rapidly absorbed by both anatomists and philosophers.

Scheiner stated that he had dissected the eyes of many animals, but had not had the opportunity of examining a human eye. Later developments in medicine, in which human eyes were available for study, confirmed the general scheme for the vertebrate eye illustrated by Scheiner. The gross anatomy of the eye as illustrated by Helmholtz is shown in Figure 4.5c. Scheiner used images reflected from the surface of the cornea in order to estimate its curvature, but Helmholtz

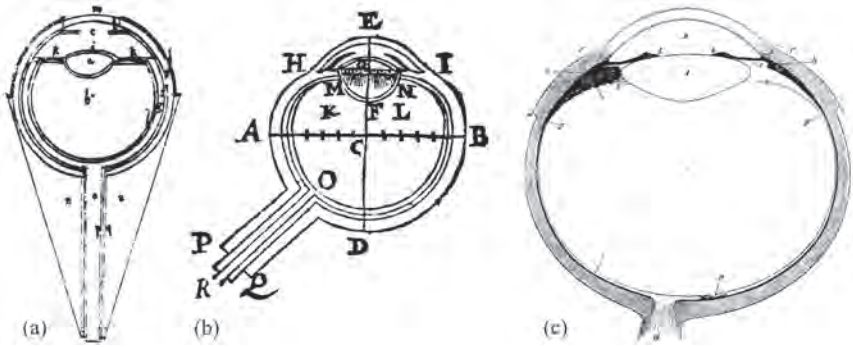


Figure 4.5 Diagrams of the eye from (a) Platter (1583), (b) Scheiner (1619) and (c) Helmholtz (1867).

was able to employ a more sophisticated range of optical instruments, like the ophthalmometer and the ophthalmoscope (which he invented), to determine the detailed structure of the eye. Since Helmholtz understood the mechanism of accommodation, the suspensory ligaments are represented in more detail than was the case for Scheiner's diagram.

A modern schematic eye is shown in Figure 4.6. Light first strikes the curved, transparent outer layer of the eye, called the cornea. If the light is not perpendicular to the cornea there is a change in the path of light rays (refraction), because the cornea has an optical density higher than that of air, and the velocity of light is reduced. The curvature of the cornea causes the rays of light from an object which

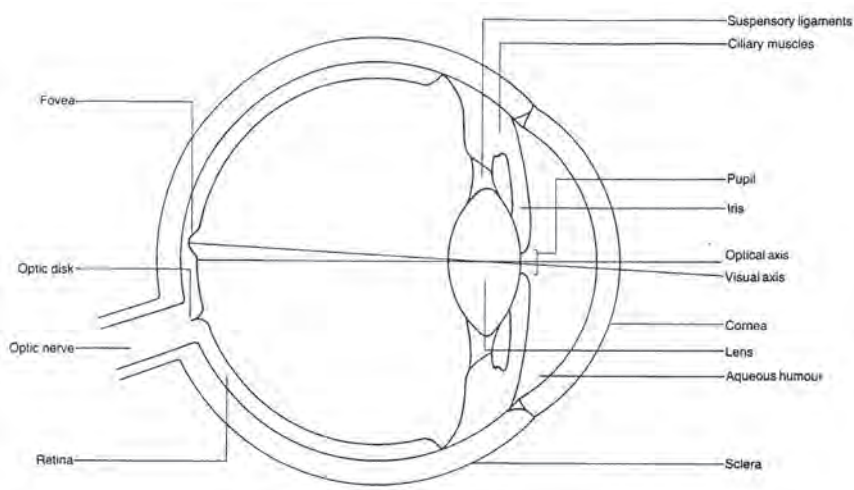


Figure 4.6 Schematic diagram of the vertebrate eye. In reality the eye is not normally a perfect sphere, but is elongated on the horizontal axis. Most of the focusing of light occurs at the curved boundary between air and the cornea. The image formed on the retina is inverted and left–right reversed.

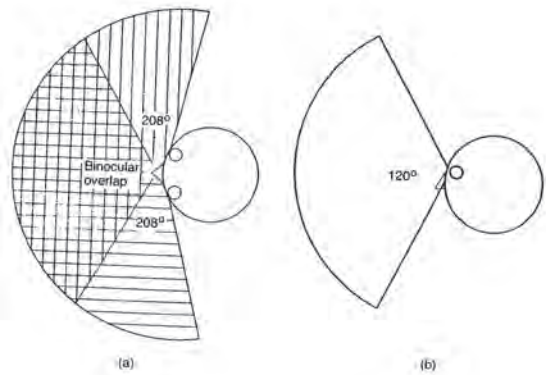


Figure 4.7 The horizontal (a) and vertical (b) field of view. The hatched area of the horizontal field represents the area of binocular overlap between the fields of the two eyes.

strikes it at varying points to be refracted by varying amounts, so that they are brought to a focus inside the eye, close to the plane of the retina. The curvature of the cornea also allows it to refract light from directions somewhat behind the observer. The full field of view extends through about 208 deg horizontally for a forward-pointing eye, although light from this region is imaged on the extreme periphery of the retina, and is blocked by the nose and head on one side. Vertically, the field of view is about 120 deg (Figure 4.7). Note that these values define the range within which light is imaged on some part of the retina; whether or not anything is seen depends on processes that occur after the absorption of light energy by the retina. It is the difference between the optical density of the cornea and air which causes refraction, which is why vision underwater is usually less distinct; there is very little refraction at the boundary between water and the cornea so that little focusing occurs.

In fact almost all the focusing power of the eye resides in the cornea, rather than in the lens. The lens provides an adjustable fine focusing power, called accommodation, due to the fact that its shape can be changed. Its normal shape is approximately spherical, but this is changed by the application of tension by ligaments attached to the transparent sack in which the lens is contained. This tension causes the lens to become more elongated, and to apply less change to the path of light rays. The ligaments themselves are controlled by a circular ring of muscles (the ciliary muscles) which when contracted allow the ligaments to reduce the tension they apply, and the lens to assume its normal spherical shape. In its most elongated state, the lens in a normal eye will bring rays from an object at optical infinity (defined as a distance greater than 6 m or about 20 feet) to a focus on the retina. This is referred to as the far point of accommodation. At its most nearly spherical, it will focus rays from an object at a distance of about 15 cm from the eye (the near point). An emmetropic, or normal, eye is capable of adjusting accommodation within the range between these far and near points. Emmetropia is the result of the focusing power of the lens combined with the

length of the eye, either of which may be inappropriate, with results that are described in the next section.

When someone's lens is removed, usually due to the presence of a cataract which makes it opaque, then vision is generally little impaired, provided the illumination is high. This is due to focusing by the cornea, and to the presence of a variable aperture in the eye's optical system, called the pupil. The pupil is created by a ring of muscles (the iris), which can expand and contract to adjust the size of the pupil, and thus the amount of light entering the eye. In high light levels, the pupil contracts to a minimum diameter (about 1 mm), and it functions like the aperture in a camera obscura, by assisting in the focusing of light. More precisely, a small pupil diameter increases the depth of focus of the eye, which is the range of object distances for which a clear image can be obtained without adjustment of the lens. By maximising the depth of focus with a small pupil, someone without a lens can see clearly over a reasonable range of distances, but this is reduced in lower light levels. In addition, there is a surprising side-effect: the lens normally absorbs most of the ultraviolet radiation that strikes the eye, and if it is removed, people report being able to see deeper shades of blue beyond the limit of visibility for normal observers. Pupil diameter can also vary for reasons other than the ambient light level. It is influenced by attention, and dilates when we try to carry out a difficult or important task. Emotions can also have an effect, as the pupil will also dilate with heightened arousal, for example due to fear. The drug atropine, which was used to cause maximum pupil dilation before clinical examination of the eye, takes its name from the plant *Atropa Belladonna* (Deadly Nightshade) from which it can be extracted. The term Belladonna means 'beautiful woman' and in the past it was not uncommon for women to use eye drops made from the plant in order to create pupil dilation, and perhaps also an increase in attractiveness. Any beneficial effects may have been due to the implication of attentive interest conveyed by the size of the pupils.

Limitations of optical performance

This sampling of the light field is the function of the image-forming components of the eye. The simplest image-forming system is an enclosed hollow space, with a small aperture through which light can enter. Any device like this will form an image on the internal surface opposite the aperture, and is called a camera obscura (Figure 4.1). A sample of the light field enters the camera obscura through the aperture, and if this is small then only a small fraction of the total light energy will be able to enter. However, the rays of light will be limited to those which form an image. The quality of the image (its sharpness and brightness) depends on the size of the aperture. A small aperture will give a sharp image, but it will also be faint. A camera obscura works best when the external illumination is high, and the aperture is small. By the addition of some means of focusing the light from distant objects, the image quality can be maintained with a larger aperture. Focusing consists of bringing together rays of light emitted from a single point in the environment to a single point in the image. Light is emitted in all directions

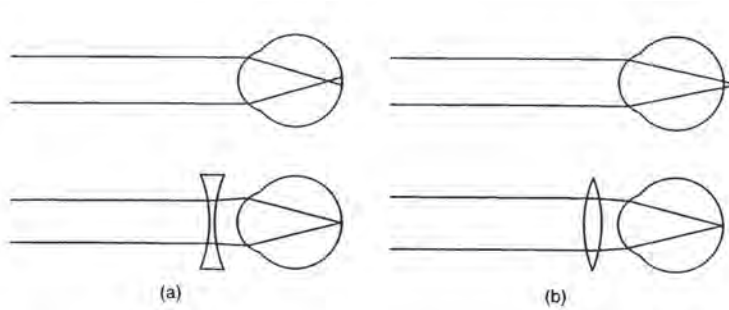


Figure 4.8 (a) Myopic and (b) hypermetropic eyes, showing the effects of optical correction. In the myopic eye, light rays from optical infinity are brought to a focus in front of the retina, which is corrected with a concave (negative) lens. The focal plane of the hypermetropic eye lies behind the retina, and this can be corrected by a convex (positive) lens.

from objects, but if the object is far enough from the imaging device then the rays which enter it will be approximately parallel, and the object is said to be at optical infinity. This distance is conventionally taken to be 6 m. If the object is close then the rays from a point on it which enter the aperture are divergent, and need greater focusing, or bending, in order to be brought to a single point in the image.

In practice, the optical functioning of the normal eye falls short of ideal in a number of ways. In addition, individuals may have various sorts of deficiency which further degrade the formation of images. Myopic (short-sighted) eyes are unable to focus rays from optical infinity on the retina, but only to a point in front of it. This is due either to the eyeball being too long, or the power of the lens being too great, or a combination of the two factors (see Figure 4.8). The far point at which myopic vision is distinct is close enough for rays from an object to be divergent, and the near point is also closer than it is for emmetropic eyes. Hypermetropic (long-sighted) eyes are the converse case; rays from optical infinity can be focused, but not those from nearer objects, and the near point is further away. Both these states can be corrected by suitable lenses; myopia with a negative (concave) lens that makes the light rays diverge as if coming from a nearer point, and hypermetropia with a positive (convex) lens that causes a corresponding convergence. In the relatively unusual event that an individual has one myopic eye, and one that is emmetropic or even hypermetropic, each can be corrected with the appropriate lens. However, the resulting images in the two eyes are of different sizes, since a convex lens magnifies the image while a concave one reduces it. This condition, known as aniseikonia, causes problems for binocular vision, since it may not be possible to fuse two such different-sized images, with the result that one is suppressed.

As ageing proceeds, it becomes harder to change the accommodation of the eye, largely due to a loss of elasticity in the lens. This is known as presbyopia, and it affects most people eventually to a greater or lesser extent, even if they were emmetropic as young adults. The effect of presbyopia is most marked in

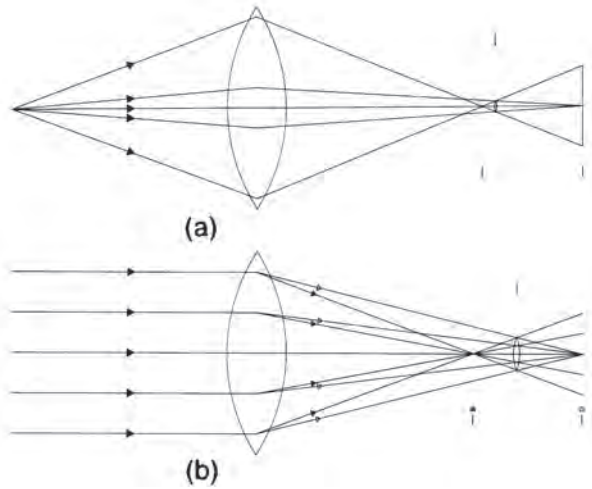


Figure 4.9 (a) Spherical aberration. The light rays striking the outer parts of the lens are refracted more than those near the axis of the lens. (b) Chromatic aberration. White light striking a lens will result in different refraction of the spectral components, so that short wavelengths are refracted most and focus closer to the lens than long wavelengths. In both illustrations the ellipses denote the disc of least confusion – the blur circle with the smallest diameter.

people who were hypermetropic, since it further restricts the range between the near and far points. By contrast, a myopic person will come closer to emmetropia as presbyopia sets in. When visual perception is impaired by presbyopia, it may be some time before the individual is aware of the extent of the impairment, or before he or she is prepared to admit to it, if good eyesight has been a source of pride. This may cause problems particularly in the work place, as errors and accidents may become more probable.

Even if the eye is emmetropic, or has been corrected by spectacles, the image formed on the retina is subject to a number of distortions. The sharpness of the image is reduced away from the centre of vision, which is a point on the retina approximately in line with the centre of the lens; it is therefore traversed by light rays without appreciable refraction. It is not surprising that the highest concentration of light-sensitive cells is found at this point. The best image is still appreciably blurred, and light energy from a point in the environment is spread out across an area, called the blur circle, rather than being concentrated at a point on the retina. This is due to spherical aberration in which light falling on the outer parts of the lens is refracted more than that near the centre (Figure 4.9a). Furthermore, different wavelengths are not focused equally; typically, the focal plane for red light lies somewhat further from the lens than that for blue light, so that if the eye accommodates to bring one wavelength into focus, others are blurred. This difference in focusing power for different wavelengths is called chromatic aberration, and it is quite marked in the human eye (Figure 4.9b). With white light, only the middle (yellow) wavelengths are usually in good

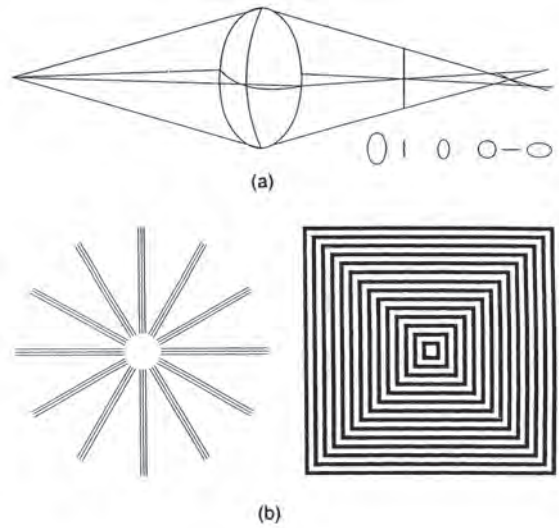


Figure 4.10 (a) Focusing of light by an astigmatic lens. The lens shown is more curved in the horizontal than in the vertical axis. Rays from a point cannot be brought to a single point of focus. Instead there are two focal planes, corresponding to the two axes of curvature. At intermediate positions, as shown, a point source is represented by an elliptical distribution of light. Astigmatism can be corrected with a lens whose curvatures are opposite to those of the lens. (b) The effect of astigmatism on the appearance of lines in different orientations. If you have an uncorrected astigmatism, it will not be possible to bring all orientations into equally sharp focus at the same time. One set of lines in a given orientation will appear darker than the rest. If the page is rotated, the lines concerned will change, corresponding to their orientation with respect to the eye.

focus; the shorter and longer wavelengths are focused in front of and behind the retina, respectively. This means that the image of a white disc consists of a central yellow area surrounded by red and blue coloured fringes. Nevertheless we are not generally aware of chromatic aberration, and there is evidently some process of compensation in the visual system which removes its effects.

Spherical and chromatic aberrations occur even if the transparent surfaces of the eye are parts of perfect spheres. Such symmetry is rare in biological systems, and the eyes usually have optical surfaces that are not spherical. The most common form for the cornea is that the curvature differs in directions at right angles to one another. This has the effect of producing two focal points for the eye, corresponding to each of the curvatures (see Figures 4.9 and 4.10) and is known as astigmatism. Thus if a horizontal line is in focus, a vertical one will not be, and vice versa. When this difference is large enough to interfere with vision, it can be corrected with a cylindrical lens that is shaped so as to focus one orientation more strongly than another, and thus compensate for the deficiency of the eye.

In addition to these distortions imposed by the optics of the eye, various effects are produced by the media through which the light passes. The cornea, lens, and fluids inside the eye are not perfectly transparent, and alter the spectral composition

of the light. In particular, the lens absorbs long and short wavelengths more than middle ones, giving it a yellowish appearance. This yellowing increases with age, leading to shifts in the appearance of colours. The fluids in the eye may develop small areas of opacity, where light is absorbed more strongly. These appear to an observer as small dark blobs, called muscae volitantes or floaters, which seem to move about as the eyes do. If the basis for the floaters is not understood, their visibility may be attributed to objects in the outside world, which seem to be moving with very high velocities; some of the reported sightings of UFOs have been traced to this source. After the light has passed through the ocular media it must pass through the layer of blood vessels and nerve fibres that lie above them before it strikes the sensitive elements in the retina. This seemingly inefficient arrangement is probably a consequence of biological constraints on the evolution of the retina as an outgrowth of the brain, and it means that light is both absorbed and scattered before it can be detected. You can in fact see the blood vessels in the eye, by shining a small torch into the corner of the eye while in an otherwise dark room. It helps to close the eye partially and to move the torch around. It usually takes some practice to produce the effect for the first time, but it is very striking once seen. The blood vessels appear as a pattern of branches, like a tree, that appears and disappears as the torch is moved around. Following Purkinje's description of their visibility (see Figure 2.30), the effect is referred to as the Purkinje tree. In fact you are seeing the shadows of the vessels, since less light reaches the retina behind where they lie. These and other sorts of perceptions produced by the structure of the eye itself are known as entoptic phenomena.

Since such effort is applied to the correction of individual errors in the optics of the eye, it might be thought that the quality of the image was of paramount importance to perception. In fact, most of the distortions of the image pass unnoticed, and it is remarkable how effectively we perceive despite their presence. Why then is it important for the image to be at least approximately in focus? The eye is a means of obtaining information about the outside world, and vision depends on the information available for processing. A blurred edge carries less information, since its location is less precisely specified, and although we can recognise objects when visual information is degraded, the process is slower and more prone to error. The quality of the retinal image imposes constraints on visual performance, and acts as a limiting factor on what we can see.

Measures of optical performance

In the previous sections we have examined the principles of image formation by the eye, and identified some of the common faults which can alter or reduce the information available for perception. An important area of study in vision is concerned with measuring the performance of the eye. This is useful for various purposes, such as identifying and correcting anomalies, and finding out the limits of normal vision in applied settings. Assessing visual resolution is an ancient practice, and much was learned from the techniques employed. The study of astronomy assisted the understanding of vision in many ways, not least

concerning the limits of visual acuity. Classical considerations of acuity were in terms of distance or distant objects. Distinguishing double stars provided one of the earliest forms of assessing the acuity of vision, and it was so applied in Egypt over 5,000 years ago. Thus, viewing distant objects, like stars, also played a role in determining the limits of spatial resolution and the variations that occur within the visual field. Euclid in his *Optics* noted that rectangular objects appeared rounded at a distance. Ptolemy provided a general statement regarding the greater distinctiveness of vision along the axis of the eye as opposed to more lateral rays. He accounted for this by saying that perpendicular rays were stronger than oblique ones. An alternative technique, introduced in the seventeenth century by Benito Daça de Valdes (1591–1634), involved presenting small objects of regular size and determining the distance at which they could not be discerned. For example, he measured the distance at which a row of mustard seeds could not be counted; he also varied the distance at which small print could be read. These measures were used to prescribe the strength of eye glasses required.

An empirical demonstration of the variations of visual resolution in central and peripheral vision was provided in the eleventh century by Ibn al-Haytham. He used a board to present two equivalently written words in central and peripheral locations; the central word was more easily read than the lateral one. The indistinctness of the lateral word increased with its movement into the periphery. The experiment was conducted with monocular and binocular observation with essentially the same results. He moved the written words peripherally until the letters became illegible. The magnitudes of these extents were not given, and many centuries were to pass before values were derived. The small region of high visual resolution had long been known, and from the eighteenth century it provided the stimulus for examining eye movements.

In the seventeenth century, Robert Hooke determined the limits of visual resolution to be around 1/60th of a degree (1 minute of arc) for most individuals, but variations, depending upon stimulus and the conditions of stimulation were also established by later researchers. Differences between using one or two lines were noted, as was the effect of stimulus intensity. There were even attempts to vary the sizes of the stimuli. A great deal was learned about visual resolution by this simple procedure of varying the distance of observation, but it did not transfer to clinical studies.

Herman Snellen (Figure 4.11) transformed the manner in which visual resolution is assessed. In contrast to previous methods of measuring acuity, which varied the distance of viewing a constant stimulus, he varied the sizes of stimuli. He developed a test, using letters of different dimensions, that was and still is widely used by opticians to determine the strength of optical corrections for short and long sightedness. Snellen's isolated shapes were called optotypes, and a modification of them, avoiding rectangular elements, was introduced by Edmond Landolt (Figure 4.11) in 1888. The Landolt C or ring consisted of a circle with a gap in it; the orientation of the gap could be varied, and the observer's task was to detect the orientation of the gap in the ring. After some competition between

Figure 4.11 Left, **Snellen's letters** A barely discernible portrait of Herman Snellen (1834–1908) is presented within letters from a modified chart for determining visual acuity. Right, **Landolt Cs** The portrait of Edmond Landolt (1846–1926) is combined with 'acuity' written in circular letters, the C of which has its opening right, above, left, and below. It has been said that the history of visual acuity is littered with still-born charts. Those of Snellen and Landolt continue to display rude health.

Snellen's letters and Landolt's C, the latter became the favoured stimulus for laboratory studies, and Snellen's letter chart was adopted by clinicians.

Snellen's letters are familiar to most people as the result of having eye-tests. Such tests employ a type of psychophysics, modified to take account of the problems of taking measurements from a cross-section of the population. It is important to realise that measurements are usually of perceptual rather than optical performance, in that people are asked to say what they can see when looking at standardised visual patterns. Under these circumstances the response may be influenced by other factors, since even the most skilled observer cannot simply report the events that take place on the retina. Nevertheless, careful use of psychophysical methods can ensure that subjective reports are a useful indication of the quality of the information available to the eye. While most of our knowledge is obtained in this way, it is also possible to make objective assessments using instruments such as the ophthalmoscope. If the subjective and objective measures differ, then it can be taken that this is due to the processing of information subsequent to the eye. For example, there is a condition known as amblyopia, in which an individual has blurred vision in one eye which cannot be corrected by any optical means. Therefore, the poor acuity is not due to the optics of the eye, but to the ways the neural signals are processed in the visual system. Amblyopia is often a consequence of an uncorrected squint during childhood, leading to failure of normal binocular vision, and the suppression of information from one eye.

The standard of vision is most commonly expressed in terms of acuity, which is, at the simplest, the ability to discriminate detail. Suppose that there are two points of light on a dark background. As described in the previous section, each will be imaged on the retina with a degree of blur, even in an emmetropic eye. More precisely, this means that the light energy will be spread out over some area,

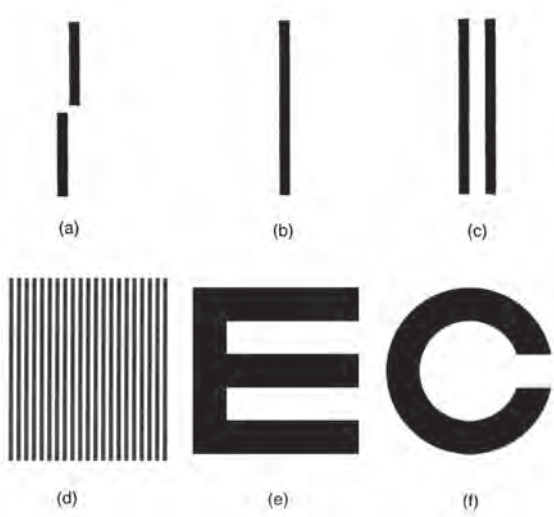


Figure 4.12 Examples of patterns used to test visual acuity. In each case the test is based on establishing the limiting conditions for discriminating a designated feature. (a) Vernier acuity. Detection of misalignment between the two line segments. This is commonly found to be very much better than would be expected from other acuity measures. (b) Detection of the presence of a single line. This depends upon the length of the line, and its orientation. (c) Two-line discrimination. The minimum separation is measured at which two lines are seen, rather than one. (d) Grating acuity. The minimum contrast at which a grating with a given spacing between the lines can be discriminated from a uniform grey field. (e) Snellen letter and (f) Landolt C or ring, as commonly used for ophthalmic assessments. The Snellen letters are read from a chart on which they appear with decreasing size. The Landolt C is shown in different orientations, with gaps of decreasing size, until a point is reached where it cannot be discriminated from a circle with a continuous circumference.

and a graph of its distribution would show a central peak with a tailing off at each side. The greater the blur, the more spread out this distribution will be. If the two points are physically separated by a sufficient distance, their two distributions will not overlap; that is, their representations on the retina will also be physically distinct, and are said to be optically resolved. As they are brought closer together, the two distributions of light energy will become increasingly combined, until they merge into one If we ask an observer to indicate the minimum separation at which the points appear separate, then we have a measure of visual acuity. If the observer's eye needs optical correction, then we would expect to find that the minimum separation is larger than that for other people. Clearly, this presupposes some accepted standard of performance against which someone can be assessed, and a standardised test pattern at which to look.

Snellen's letters and Landolt's C are not the only patterns that have been proposed, and others are useful under different circumstances (Figure 4.12). The familiar opticians' Snellen chart uses letters of the alphabet of decreasing size, and tests the ability to discriminate the separate features of each letter well enough

to identify them. While this has some relevance to real-life visual tasks, it has the drawback that we have little knowledge of the processes that underlie the recognition of letters. The Landolt C chart presents the ring in various sizes and orientations; the observer's task is to identify the location of the gap, whether up, down, left, or right. As the 'C' is made progressively smaller, there comes a point at which performance in detecting the gap is no better than chance, and this is the limit of acuity. In general, visual acuity is expressed in terms of the minimum resolvable visual angle, for which the normal standard is a gap subtending 1 minute of arc. Visual angles (the angle subtended at the eye) are measured in degrees, and an approximate indication of 1 degree is the angle subtended by the index finger nail when the arm is fully extended. Acuity is stated as a fraction, relating a standard testing distance (6 m or 20 feet) to the distance at which a gap would subtend 1 minute of arc. Thus an acuity of 6/6 (which is the same as 20/20) means that performance is normal, whereas 6/3 means that a gap could be resolved at 6 metres which would subtend 1 min at 3 metres, i.e. a better than normal acuity. Similarly, 6/18 means that the minimum gap that can be seen subtends 3 min at 6 metres, and performance is a third of normal. While these sorts of test have considerable practical utility, they have a number of drawbacks for scientific research on perception. Even if they are carried out under carefully controlled conditions of illumination, the fact remains that both Snellen letters and Landolt C's are complex patterns, from which it is not easy to generalise.

One approach to specifying visual performance in the most general terms derives from measurements of acuity for grating patterns. A grating consists of a number of light and dark lines with a particular width and orientation. Grating acuity is assessed by reducing the width of the light and dark lines until the grating can just be discriminated from an unpatterned grey field of the same average luminance. If the boundary between the light and dark lines is sharp and they are equally wide, then the grating is said to have a square-wave profile. This can be seen in Figure 4.13, which shows a graph of luminance in a square wave grating across the width of the field. A different luminance profile is produced if the change from light to dark is a gradual one. An example of this is the sine wave grating, in which luminance varies across the test field according to a sine function (see Figure 4.14). A cosine grating would have a similar profile, but shifted so that the dark lines coincide with the middle luminance of a sine wave grating. In this case the difference is in the phase of the gratings, that is, in the relative location of the maxima and minima in luminance. Phase is expressed in terms of angles, and a cosine grating is 90 degrees out of phase with a sine wave one. The importance of sine wave gratings lies in their use as general tests of the performance of an optical system. To understand the nature and significance of this work, it is necessary to describe some basic concepts in the specification of signals, whether optical, electronic, or acoustic.

In the early nineteenth century, the French mathematician Jean Fourier (1768–1830) proposed the theorem that any complex change in the state of a system could be described in terms of the combination of simple sine wave oscillations of different frequency and amplitude. Thus a pure tone is a single frequency of

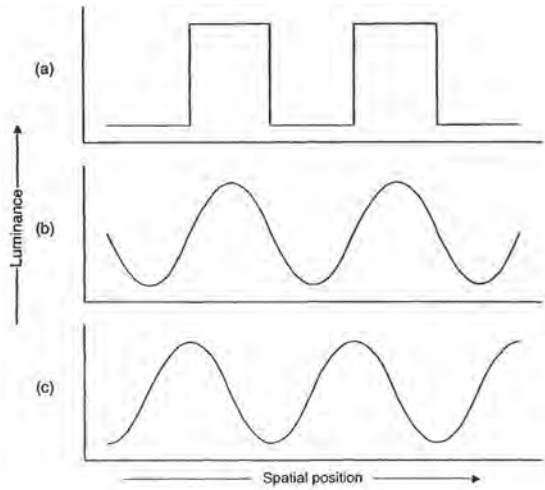


Figure 4.13 Illustration of the way in which luminance varies with spatial position for various spatial waveforms. In (a), a square wave, there are sharp transitions from light to dark, forming a series of bars with sharp edges. In (b) a sine wave and (c) a cosine wave, the transitions are more gradual. Sine waves differ from cosine waves only in the location of the maxima and minima, i.e., in phase. A waveform is fully described by its frequency (the number of oscillations in a given interval), its amplitude (the size of the oscillations), and its phase. Any complex waveform can be created from a combination of sine waves with appropriate amplitudes, phases and frequencies.

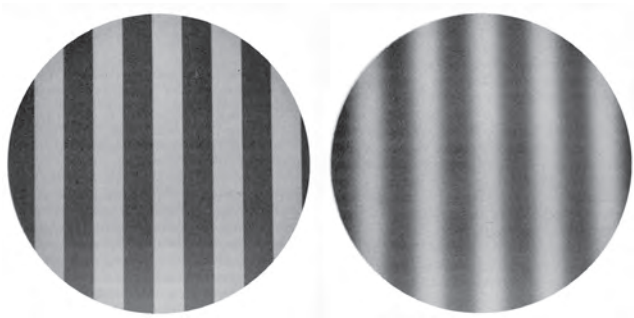


Figure 4.14 A square wave and sine wave grating. Note the gradual variation in brightness of the sine wave at the transition from light to dark.

vibration, whose amplitude varies sinusoidally over time. Any complex sound can be described in terms of a combination of pure tones, with suitably chosen frequencies and amplitudes. This provides a very powerful means of specifying sounds, and of testing the response to them by physical systems. By testing the response with a range of pure tones, it is possible to predict performance with any arbitrary sound that might occur. This can be applied, for example, to the specification of performance in loud-speaker systems.

In order to apply these concepts to vision, it is necessary to think in terms of spatial rather than temporal frequencies; that is, rather than measuring the number of cycles of sound pressure in one second the index is the number of cycles of luminance in one degree of visual angle. Visual stimuli have the property of being extended over space, and intensity can vary in this dimension also. Thus a static pattern which changes from light to dark according to a sine wave function constitutes a spatial, rather than temporal, frequency (Figure 4.14).

Any optical system can be tested with a range of such spatial frequencies in order to find out how well they are transmitted. The relevant measure is the contrast of the test pattern emerging from the optical system, compared with that which is fed in. Loss of contrast shows up as a reduction in the difference between the light and dark areas. The overall optical performance is expressed as a modulation transfer function (MTF), which is the relationship between contrast transmission and spatial frequency. Direct measurements of the MTF for the human eye might be made by measuring the contrast of gratings projected onto the retinal surface, for example with an ophthalmoscope and a photodetector. More simply, an observer can be asked to report when the contrast of a sine wave grating with a certain spatial frequency is sufficient for it to be just visible. In this case the measurement is of the MTF of the observer, treated as a single system, and it incorporates the effects of neural as well as optical processes. Such an MTF is called a contrast sensitivity function (CSF), and an example is shown in Figure 4.15.

The human CSF clearly shows a differential response to different spatial frequencies. For example, there is a rapid drop in response above about 12

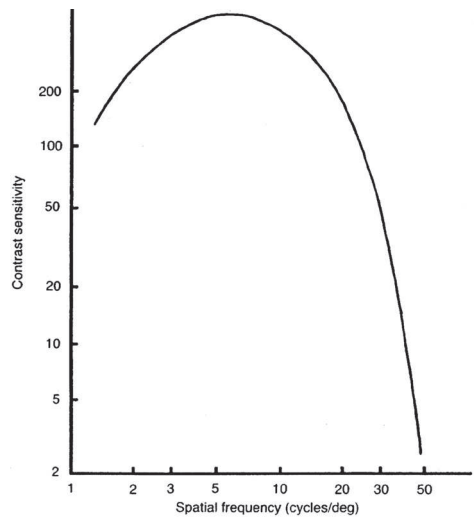


Figure 4.15 A typical human contrast sensitivity function. Contrast is measured by the ratio of the difference between the luminances of lightest and darkest parts of the grating, divided by their sum. Contrast sensitivity is the reciprocal of the contrast at which a grating of a given spatial frequency is just discriminable from a uniform grey field. Peak sensitivity occurs at around 3–5 cycles/deg.

cycles per degree, and a more gradual decline below about 6 cycles per degree. The high frequency cut-off is probably due to the optical characteristics of the eye, and is related to the diameter of the pupil. It is in fact very close to the ideal performance of an optical system with the dimensions of the eye. The low frequency cut-off is more influenced by signal processing later in the visual system. Note that the scale of spatial frequency is logarithmic, which can give a somewhat misleading impression of the shape of the function. The CSF is an MTF for gratings at threshold contrast, and other techniques can be used to find MTFs for supra-threshold gratings, although this is more difficult. The application of Fourier analysis, and the use of visual stimuli in the form of sinusoidal gratings with a defined spatial frequency, has led to the development of theories regarding the processing of visual information by the brain. Another consequence of this research has been the development of new tests of visual resolution which vary contrast as well as spatial frequency.

Measures of optical performance of the eye have in the past confounded two of its characteristics – the optics of the eye and the neural characteristics of the retina and its connections. A recent development, called adaptive optics, has enabled the separation of the eye's optics from its neural responses. The technique is used in astronomy to remove aberrations like atmospheric distortions and it can similarly remove aberrations of the human eye, like astigmatism. With the benefit of adaptive optical procedures, far higher spatial frequencies can be detected than under conditions of normal observation.

Ocular microanatomy

Study of the microscopic structure of the eye was only possible after the invention of suitable optical instruments for magnification. Cells were described soon after the first microscopes were focused on organic matter in the seventeenth century. Robert Hooke, in his book *Micrographia* published in 1665, gave them this name when he observed the structure of sections of cork. Antonius van Leeuwenhoek (1632–1723) gave accounts of several animal cells, including nerve fibres, when he directed his simple magnifier to animal tissue. He had neither the benefit of achromatic lenses nor staining methods when he observed the structure of optic nerves; he did, however, have a sharp knife to cut sections of the nerve. Leeuwenhoek was attempting to discover whether the nerves were hollow, as had anciently been believed, but he concluded that they were not. Relatively little attention was paid to these observations because the resolving powers of the microscopes were poor, due to the aberrations introduced by their optical components. The microscopic world was transformed by the introduction of powerful achromatic instruments in the 1820s, and rapid advances were made thereafter. The cell doctrine was most clearly articulated at the end of the next decade but it was not until the end of the nineteenth century that the neuron doctrine began to be widely accepted.

Anatomical knowledge increased in the eighteenth and particularly in the nineteenth centuries – the power of the microscope was brought to bear on

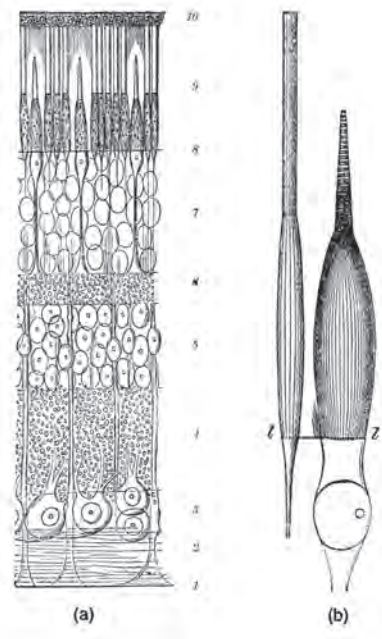


Figure 4.16 (a) Diagram of the microanatomy of the retina from Schultze (1866), who distinguished between the rods and cones (b) and assigned different functions to them.

unravelling the detailed structure of the retina and the nerve pathways, and structure was in turn related to function. For example, in the 1860s two different types of light-sensitive cells were found in the retina; they were called rods and cones because of their shapes when observed under the microscope (Figure 4.16). It was noted that the proportion of rods and cones varied over species. Indeed, some species only had one type or the other, and in all cases this could be related to the light conditions in which they were active. The retinas of nocturnal animals had rods but few or no cones, whereas those active only in daylight had cones but few or no rods: humans, who are active and can see under both natural light conditions, have a mixture of rods and cones. Accordingly, it was proposed that rods and cones had differing sensitivities and served different functions. Rods are able to detect light at lower intensities than cones, but cones are involved in colour vision and rods are not. Both rods and cones require appropriate pathways to the brain in order for us to experience light; however, in the nineteenth century relatively little was known about pathways in the brain, and so physiological analyses of vision tended to be restricted to the structures in the eye.

The spirit of these exciting times was reflected in what is perhaps the greatest book written on vision, the *Treatise on physiological optics* by Helmholtz. It was originally published in three separate volumes during the 1850s and 1860s, and the complete treatise was translated from German into English in the early twentieth century to mark the centenary of Helmholtz's birth. Helmholtz was both a physicist

and a physiologist, and he amalgamated the strengths of these two disciplines in the analysis of vision. We have mentioned several of his contributions to visual science in earlier chapters: the elucidation of the mechanism of accommodation, the championing of a theory of colour vision based on three colour channels, and the invention of the ophthalmoscope for examining the inside of the eye. The ophthalmoscope is an instrument which directs light into the eye and light reflected from the retina and related structures is magnified. It opened a new world to ophthalmologists and many new discoveries were made from the middle of the nineteenth century.

Amongst these was the colour of the retina itself. Helmholtz remarked: "The ground of the eye looks reddish, and indeed first of all round about the optic nerve of a somewhat clear, light-red, the darker, on the contrary, the farther you pass from that place". In the same year (1851), direct observation of the retinas of frogs by Heinrich Müller (1820–1864) also led to descriptions of its reddish colour. Müller's discovery of the coloured retinal receptors was not actively pursued for another two decades when his pioneering observation was extended by Franz Christian Boll (1849–1879). He provided convincing evidence that the red appearance of the retina reflected the colour of the rod pigment and that it was bleached by light. He speculated that this was the basis for the reddish appearance of the human retina seen through an ophthalmoscope.

The pigment in the rods is now referred to as visual purple rather than visual red, as Boll had called it. This is due in large measure to the researches of Willy Kühne (1837–1900) who was able to extract rhodopsin from the rods of frogs and rabbits, and his elegant experiments provided the foundations upon which subsequent studies were built. He showed that the rate of bleaching was dependent not only on the intensity of light but also on its wavelength. The visual purple was confined to rods and was not seen in the foveas of humans. Most significantly, Kühne established the 'visual cycle': visual purple in the rods is bleached by light to form visual yellow which in turn is transformed to visual white. Over half a century later, George Wald (1906–1997) fashioned some chemical flesh on Kühne's cyclic skeleton. Initially working with frog retinas, Wald examined the chemical reactions when it was in the purple, yellow and white stages. He established that light converts rhodopsin to retinal and opsin through a series of intermediate reactions involving metarhodopsins; one of the intermediates results in excitation of the photoreceptor cell whereas another releases opsin. Retinal is reduced to vitamin A and regenerates rhodopsin, as does opsin.

Receptors

The concept that there were cells specialised to be receptors arrived relatively late on the sensory scene. When Descartes was discussing the eye he considered that the retina consisted of nerve endings and that their dimensions defined the limits of visual resolution. As a great variety of cells became visible under the microscope, those that were isolated in sense organs were called receptors and they could be related to the stimuli that excited them. Those located in well-

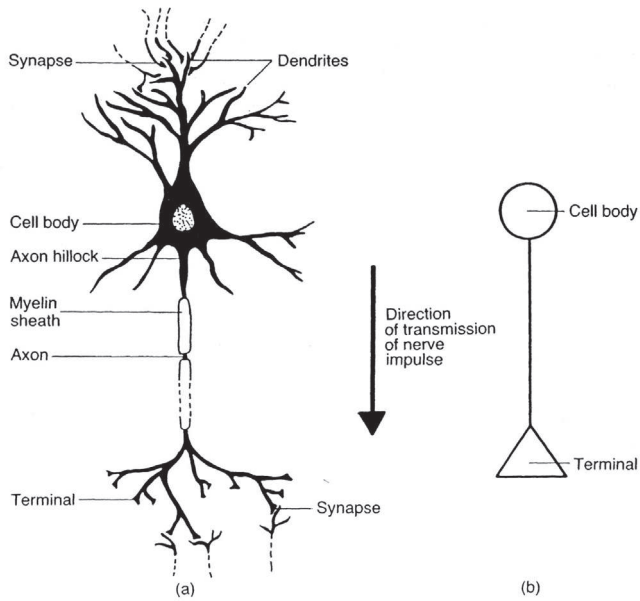


Figure 4.17 The main structures in a typical nerve cell. (a) The dendrites receive inputs, transmitted chemically across the synapses, from many other nerve cells; inputs can be excitatory and inhibitory. If the net activity passes some threshold then an action potential or nerve impulse is initiated in the axon hillock and it is transmitted along the axon to the terminals of the neuron, thereby influencing other neurons. The transmission of the action potential along the axon is facilitated by the myelin sheath, which acts as an insulator. (b) A schematic nerve cell of the type that will be used to indicate the direction of neural transmission in some of the following illustrations.

defined sense organs were named on the basis of their morphology (rods, cones, hair cells, etc.), whereas the receptors in or beneath the skin were generally named after those who first described them (e.g., Golgi tendon organs, Krause end bulbs, Meissner corpuscles, Merkel discs, Pacinian corpuscles, and Ruffini cylinders). The isolation of receptors that were specialised to respond to specific forms of environmental energy was adopted as a criterion for defining the senses.

All sensory systems function by transducing some type of environmental energy into a form that can be analysed by the cells in the central nervous system (CNS). The general structure of nerve cells (neurons) is shown schematically in Figure 4.17. Neurons communicate electrochemically: they transmit signals along the nerve fibres in pulses of fixed amplitude, hence the term nerve impulse but it is also called an action potential. Neurons do not make direct contacts with one another, and the activity of one neuron influences others chemically across the small gaps or synapses separating them via neurotransmitters. Receptors, like other nerve cells, have a resting potential difference between the inside and outside of the cell. Processes in the membrane surrounding the nerve cell retain

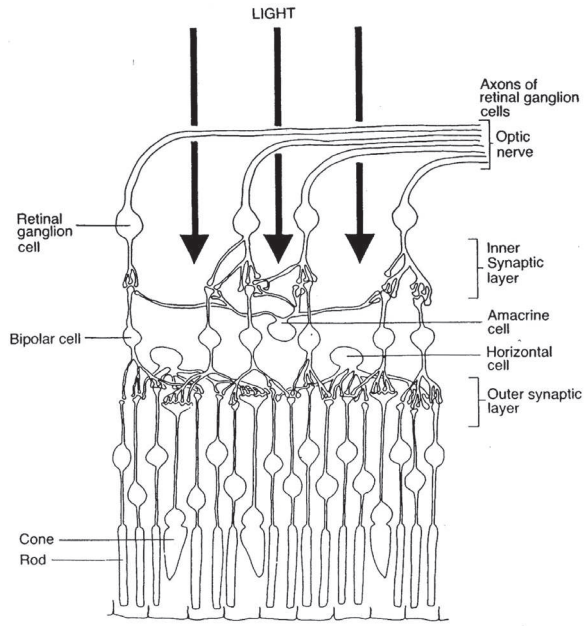


Figure 4.18 The neural structure of the retina. Light passes through the neural layers before striking the receptors (rods and cones) which contain the photosensitive pigments. The vertical organisation of the retina is from receptor to bipolar cell to retinal ganglion cell. The horizontal organisation is mediated by horizontal cells at the receptor–bipolar (outer) synaptic layer and by the amacrine cells at the bipolar–retinal ganglion cell (inner) synaptic layer.

negatively charged molecules (negative ions) within the cell so that the interior is negatively charged (by about 70 millivolts) with respect to the extracellular ions.

The transduction process usually involves a modification of the potential, so that the potential difference is reduced; this is called depolarisation. These electrical changes are graded, i.e., they will vary with the intensity of the environmental energy stimulating the receptor. The process of vision is initiated by light falling on specialised receptors in the retina (Figure 4.18). The retinal receptors contain light-sensitive pigments that are modified chemically by absorbing light. However, unlike receptors in other senses they hyperpolarise when light falls on them: the potential difference between the inside and outside of the receptor cells increases. Before describing the physiological processes at the retinal level, a little more should be said about the structure of the retina itself.

The retina is an outgrowth of the CNS and the receptors are directed towards the back of the eye. Therefore, before the light strikes the receptors it passes through the various neural structures in the retina and also the blood vessels that supply them. There are two types of receptor in the human retina, and they are called rods and cones because of their appearance under the microscope; rods have a cylindrical outer segment (which contains the photosensitive pigment

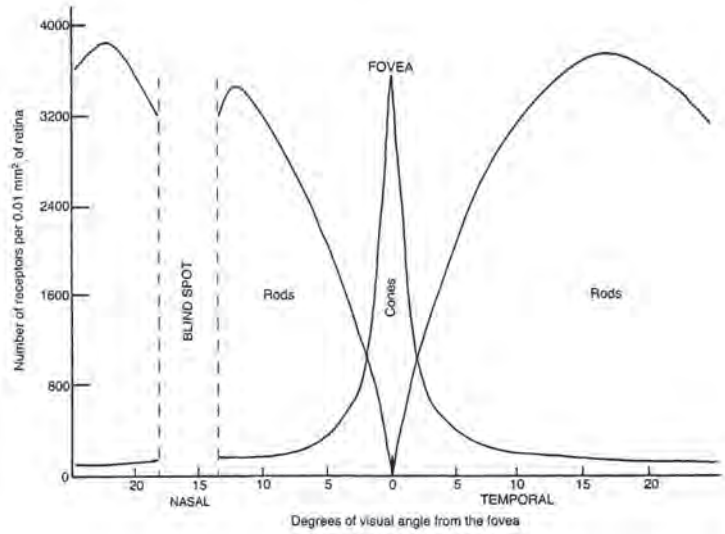


Figure 4.19 Distribution of rods and cones over a horizontal region of the retina passing through the fovea and blind spot.

molecules) and cones have tapered outer segments. The retina is estimated to consist of about 130 million receptors, the vast majority (over 120 million) of which are rods with about 6 million cones. The distributions of the rods and cones differ in a systematic way. The cones are concentrated in and around a central region of the retina, where there is a shallow depression, called the fovea. There are decreasing numbers of cones as the distance from the fovea increases, and there are none in the peripheral regions of the retina (Figure 4.19).

One part of each eye is devoid of any receptors, and it is called the blind spot (or optic disc). It is located about 17 deg towards the nasal side of each fovea, and it is the exit pathway of the nerves from the retina (the optic nerve), as well as the region where the arteries enter and veins leave the eye. We are usually unaware of this small area of blindness in each eye, because it corresponds to different parts of the visual field of each eye. Even when we use one eye alone we generally overlook this blind spot, and it is only when we employ some specific procedure, like following the instructions to observe Figure 4.20 that we can locate it.

Figure 4.20 Fixate the central cross with the right eye alone and move the book (or your head) until the black spot on the right disappears. This will happen when the spot is falling on the blind spot or optic disc, which is about 17 deg nasal of the fovea. The left spot remains visible because there are receptors at the equivalent temporal region of the retina. The left spot can be rendered invisible by fixating the cross with the left eye alone, at about the same distance.

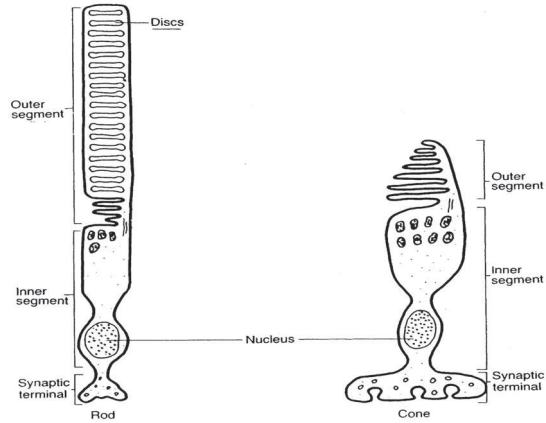


Figure 4.21 Detailed structure of rod and cone receptors. The light-sensitive pigments are located in the outer segments – stacked discs for rods and a single folded surface for cones. The inner segment is concerned with the metabolic functions of the cells.

The light-absorbing pigment molecules in the receptors are distributed in the photosensitive membrane in their outer segments (Figure 4.21). In rods the photosensitive membrane consists of separate discs stacked on top of one another, rather like a pile of coins. The photosensitive membrane in cones has a different structure: it consists of a single surface folded successively over on itself. All rods contain the same visual pigment, which is a complex protein molecule called rhodopsin. There are three different types of cones, each with a specific light-sensitive pigment. The cone pigments can absorb light throughout the visible spectrum, but one type is most sensitive to low wavelengths of light (around 420 nm), another is most sensitive to the middle region of the spectrum (around 530 nm), and the third to slightly longer wavelengths (around 565 nm).

Duplicity theory

With all these differences in mind, it is not surprising that the rods and cones operate under different conditions of light stimulation. The rods are more sensitive than the cones and respond in twilight or moonlight; rods become overloaded in daylight and are unable to function. The cones require the more intense illumination of daylight to function. This difference in sensitivity is indicated in Figure 4.22; rods are about 1,000 times more sensitive than cones. Indeed, rods are so sensitive that an individual rod can register the absorption of a single photon of light, but the activity of a single rod does not provide an impression of light. Under ideal conditions a human observer can detect a light when less than 10 photons are absorbed. The increased sensitivity of rods is combined with a slow response to light; it is estimated that a rod receptor takes about 0.3 s to signal the absorption of a photon. The cones require more light to function, but their response is about four times faster than rods. There is, therefore, a trade-off between sensitivity and

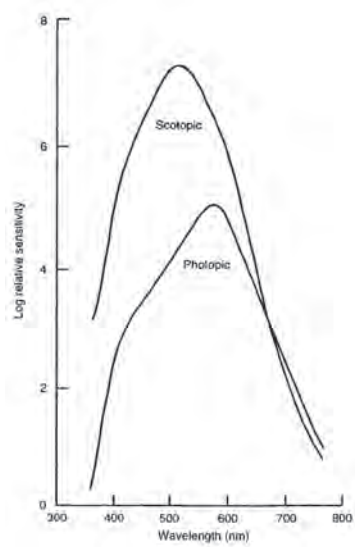


Figure 4.22 Spectral sensitivity of the rods (scotopic curve) and cones (photopic curve). The rods can detect light at lower intensities than the cones, and their peak sensitivity is for shorter wavelengths of light.

response time in rods and cones. Figure 4.22 also illustrates another aspect of rod and cone function: in dim light rods respond to the whole visible spectrum but they are most sensitive to wavelengths around 505 nm; the spectral sensitivity under bright conditions is somewhat different, showing a peak at around 555 nm. This could be the basis for some differences in our sensitivities to colours under dim or bright light conditions. Bluish-green objects appear to be brighter in dim light than they do in daylight, and conversely greenish-yellow objects appear brighter in daylight than under dim conditions. This is called the Purkinje shift, after the physiologist who first described it in 1825. We are most sensitive to light of around 505 nm in low illumination (reflecting the function of rods) and around 555 nm in bright conditions (reflecting cone function).

Another feature of this distinction is the time taken for the receptors to recover from intense light adaptation. When we have been exposed to bright daylight and then enter a dark environment (like a cinema) we are initially unable to distinguish any detail, but gradually our sensitivity improves until after about 30 min we can see quite well. It is referred to as dark adaptation; it can be measured accurately by adapting the observer to bright light for a few minutes, then placing them in darkness and measuring the detection threshold for light at regular intervals. The resulting dark adaptation curve (Figure 4.23) has two components, the initial phase shows the fast recovery of the cones and the final phase the slower recovery of the rods. If the detection threshold is measured with a small red spot presented to the fovea then only the initial component will be found, because the spot would

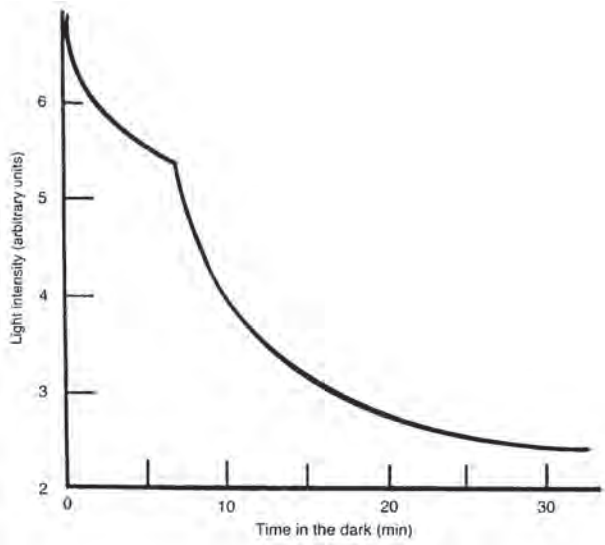


Figure 4.23 Dark adaptation or the recovery from light adaptation. The curve represents the increasing sensitivity (or decreasing thresholds) for light detection with time in the dark. The initial rapid phase reflects the recovery of the cones and the slower phase that of the rods. An asymptote is reached after about 30 min in darkness.

stimulate cones alone. Some individuals do not have any cones in the retina, and their dark adaptation curves would have the second phase alone.

Any photoreceptor can only signal the absorption of photons of light. The signals delivered will depend upon the wavelength of the light as well as its intensity, because of the differences in spectral sensitivity described above. Therefore, the same signal can result from a wide range of combinations between intensity and wavelength, and no photoreceptor can distinguish between these two dimensions. This is referred to as the principle of univariance. This principle is important when discussing differences in wavelength sensitivity between rods and cones because it applies to both of them. All rods contain the same visual pigment and so they could not discriminate on the basis of wavelength. The same is true for any given type of cone: it can only respond in one dimension, by hyperpolarising by some amount. However, there are three different types of cone, and their combined activities can mediate wavelength discrimination or colour vision. It is now considered that colour vision is based upon signals from the three cone types being combined at subsequent neural stages. This is demonstrated with regard to our ability to see colours in different regions of the visual field. Colour vision is best in and around the fovea. However, in the peripheral retina, where there are rods but no cones, we are all colour blind.

The differences in function described above are attributable to the characteristics of rods and cones, and this is referred to as duplicity theory. However, we should be cautious of treating the rods and cones alone as the bases for the effects described; the effects are the consequence of activity occurring throughout the

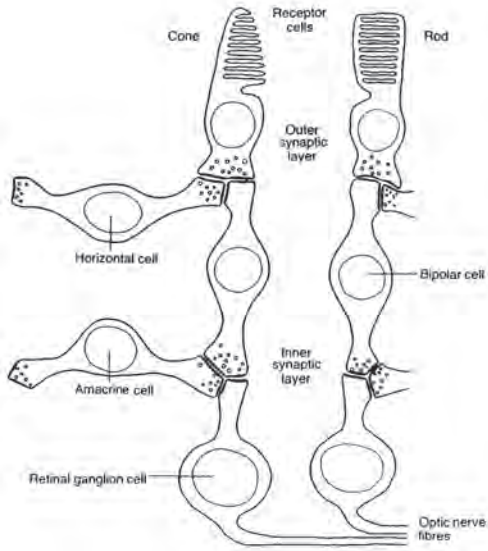


Figure 4.24 Detailed representation of the neural structures in the retina. At the outer synaptic layer the receptors can influence the activity of the bipolar and horizontal cells, and the horizontal cells can affect the receptor and bipolar cells. At the inner synaptic layer the bipolar cells can influence the activity of both the retinal ganglion and amacrine cells, and the amacrine cells can in turn affect the bipolar and retinal ganglion cells. The optic nerve is comprised of retinal ganglion cells that have been characterised in four types: midget, parasol, bistratified and biplexiform.

visual system, not just at the initial stages. The activity of the rods and cones defines the limiting light conditions under which vision is possible. Beyond these the manner in which neural information is processed will be of importance, and therefore it is necessary to examine the neural interactions within the retina and the pathways to the brain.

Retinal structure

The retina is a complex neural structure, as was evident from Figure 4.18. An even more simplified and schematic diagram of its structure is shown in Figure 4.24. The retina can be thought of as having both a vertical and a horizontal organisation: the vertical organisation corresponds to the connections leading to the brain, and these have horizontal (or lateral) connections between them. Receptors make synaptic connections with bipolar cells, which in turn make synaptic contact with retinal ganglion cells. The ganglion cells are much larger than receptors or bipolar cells; their cell bodies are in the retina, but their axons (or nerve fibres) leave the retina to form the optic nerve. There are about 1 million axons in each optic nerve, so an enormous neural convergence has taken place in the retina – from about 130 million receptors to 1 million ganglion cells. This convergence is not evenly

distributed over the retina, and is least for the foveal region. Conversely, the convergence is much greater than average in the peripheral retina. The horizontal or lateral connections occur at the first and second synaptic layers. At the first synaptic layer, between receptors and bipolar cells, there are horizontal cells, and at the second layer, between bipolar and retinal ganglion cells, there are amacrine cells. Retinal ganglion cells can be classified according to the sizes and shapes of their cell bodies. Four cell types have been classified: midget, parasol, bistratified and biplexiform, although relatively little is known about the last. In primate retina the vast majority (about 80 per cent) are midget ganglion cells; they have small cell bodies and receive their input from cones. Parasol ganglion cells make up between 8–10 per cent of the population; they have larger cell bodies and receive input from more peripheral regions of the retina in which the rods predominate. Less than 10 per cent of retinal ganglion cells are classified as bistratified; they receive input from cones. The distinction between these cell types is maintained in subsequent processing at more central sites.

The growing understanding of processing in the retina is accompanied by more detailed determination of the ways in which these chemical changes are transmitted to the brain. However, a prior problem was to trace the pathways taken from the eyes to the brain, and it is to this issue that we will now turn.

Reference notes

Porta's book on natural magic was an amalgam of mysticism, folklore, and science, but it was enormously popular. It was reprinted in many editions, and translated into several languages; the first edition (in Latin) was published in four books (1558) and it was extended to twenty books for the second edition (1589). The quotations are taken from the English translation of 1669. Porta's work is examined in some detail by Ronchi (1970). The changing ideas about the nature of light and those who changed them are described in Clegg (2001). Pirenne's books on *Vision and the eye* (1967) and *Optics, painting and photography* (1970), although now a little old, reward reading as they cover the revolution in vision, heralded by Kepler and Scheiner, with great insight and erudition. They also provide one of the clearest accounts of the camera obscura, and the second book contains many photographs taken with pinhole cameras. In addition, the relationship between the optics of cameras and eyes is discussed. In his *Popular lectures on scientific subjects*, Helmholtz (1873) remarked: "Regarded as an optical instrument, the eye is a camera obscura". The history of this much used and abused analogy has been surveyed by Wade and Finger (2001). Scheiner made explicit comparisons between art and nature (camera and eye) in his book *Rosa ursina* (1630). Scheiner's work is now more accessible in English due to the studies of Daxecker (2004).

Thomas Young's greatest contribution to science is generally considered to be in optics, and in his *Lectures* (1807/2002) he stated that "The science of optics is one of the most elegant, and the most important branches of natural and mechanical philosophy". Young extended physical, physiological, and visual optics both

experimentally and theoretically. He provided support for the wave theory of light by his experiments on interference phenomena; he described astigmatism for the first time, and he provided evidence to indicate that accommodation occurs by variation in lens curvature. Young measured the optical constants of his own eyes using an optometer of his devising, in the course of which he noted his own astigmatism. Cylindrical corrections for astigmatism were advocated soon after and charts for testing it were introduced at the end of the nineteenth century. Several of the modern tests of visual acuity were devised around that time, like Snellen's letters and the Landolt C or ring (see Wade, 2008a). Landolt's C became the favoured stimulus for laboratory studies, and Snellen's letter chart was adopted by clinicians (see Colenbrander, 2001). Modern charts, like that devised by Pelli, Robson and Wilkins (1988), incorporate variations in contrast as well as size. A history of early eye glasses, with an emphasis on the Italian contributions, has been written by Ilardi (2007). Optical disorders and their corrections are dealt with in detail by Rosenfield, Logan and Edwards (2009). Adaptive optics are applied increasingly to understand the basic optical functions of the eye, and their uses are reviewed by Roorda (2011).

The eye has been the subject of numerous books, and one entitled *The eye: A natural history* (Ings, 2007) covers many of the features presented in this chapter in a very readable way. Gregory's (1997) *Eye and brain* provides a delightful introduction to the psychology of seeing, with most of the chapters placing phenomena in their historical contexts. The books by Mather (2008) and Frisby and Stone (2010) describe the structure of the eye and retina very clearly; both books deal with the grating stimuli, providing illustrations of spatial frequency and contrast. The historical dimension is also a feature of Palmer's (1999) comprehensive treatment of *Vision science.* The anatomy and physiology of the human eye is attractively illustrated in Rogers (2011), and optical aberrations are described in Wandell (1995), as well as measurement of optical performance and human contrast sensitivity. The history of measures of visual resolution and the development of tests for acuity can be found in Wade (2004b, 2007b).

In their *Principles of neural science*, Kandel, Schwartz and Jessell (2000) describe the nature of receptors and neural transmission generally as well as giving detailed descriptions of retinal structure. Our diagrams of retinal structure have been redrawn after illustrations in their book. They also cover the initial stages of photo-transduction, and the pigments involved in it. An account of the pioneering research of Franz Boll and his competitor Willy Kühne can be read in Wade (2008c, d). Duplicity theory was proposed by Max Schultze in 1866, largely on the basis of differences in the rod and cone distributions of animals that were active in daylight or at night (see Stabell and Stabell, 2009). Coloured illustrations of the microanatomy of the retina can be seen in Yannuzzi (2010) and Besharse and Bok (2011). Details of receptor structure and distribution are given in almost all books on the eye.

5 The visual brain

The brain has been a substance of mystery for millennia, particularly since it was so difficult to examine. Functions were assigned to it on the basis of myth and magic and this continued long after its gross structure was examined by eye and dissected by knife. Alcmaeon, in the fifth century BC, described the optic nerves and proposed that they were hollow tubes. This tradition was continued by Aristotle and it was maintained for many centuries. Contrary to Aristotle, Galen (Figure 5.1), in the second century AD, believed that the origin of the visual pathways was located in the anterior ventricle of the brain, where the animal spirit could interact with the visual spirit (the pneuma), borne by the optic nerves. The optic nerves themselves came together at the optic chiasm, but each of the nerves was considered to remain on the same side (ipsilateral). That is, the visual spirit was unified in the chiasm and then sent to each eye. Galen provided an account of sight which was firmly anchored in a speculative physiology, and it was, in turn, integrated with his description of the anatomy of the eye. He also ventured, with some misgivings, into the arena of optics. In his book *On the usefulness of parts of the body* he expressed regret for introducing optical concepts in a medical text, since they were at that time deeply unfashionable. His theory of vision was physiological, and it was based on pneumatic concepts: pneuma passed along the hollow tubes of the optic nerves to interact with returning images of external

Figure 5.1 Left, Claudius Galen (c. 130–200) and right, Andreas Vesalius (1514–1564).

objects in the crystalline lens. The lens was considered to be the primary organ of vision. Galen did not rely on speculation alone; he supported the proposal that the lens was the 'seat of vision' by virtue of the blindness that results from cataracts and the sight that is restored when they are surgically removed.

Galen based his anatomy on dissections of animals, particularly monkeys, but most of his ocular anatomy was derived from dissecting the eyes of freshly slaughtered oxen. The restrictions that were subsequently placed on dissections resulted in a reliance on Greek (and particularly Galen's) works on anatomy, and they were recounted dogmatically until the time of Vesalius (Figure 5.1) over one thousand years later. From the sixteenth century, anatomy became increasingly rigorous, being based on observations from dissection, even when these were at variance with the Galenic dogma. The eye, however, was not subjected to the same detailed scrutiny as the skeleton and musculature until the seventeenth century.

Visual pathways

For centuries, diagrams of the visual pathways represented the optic chiasm but the passage of the optic nerves through it remained a matter of speculation. Vesalius repeated the separate and ipsilateral projection of the optic nerves, and it became integrated into Descartes' analysis of vision (Figure 5.2). Descartes did stress the correspondence between points on the object, those on the retina, and their projection to the brain, but it is unlikely that he was addressing the issue of corresponding points in the two retinas. His analysis of binocular vision was by the ancient analogy with a blind man holding two sticks, and it was not physiological. The union that was depicted in the pineal body reflected an attempt to match singleness of vision

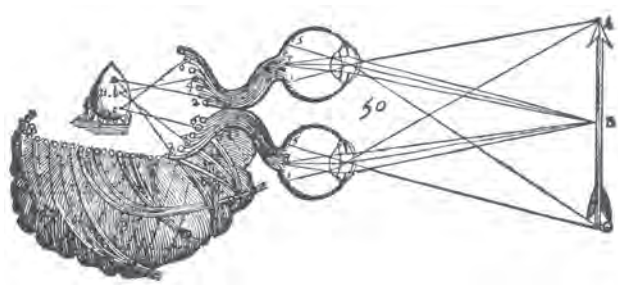


Figure 5.2 Binocular visual pathways according to Descartes: "the filaments 1–2, 3–4, 5–6, and the like compose the optic nerve and extend from the back of the eye (1, 3, 5) to the internal surface of the brain (2, 4, 6). Now assume that these threads are so arranged that if the rays that come, for example, from point A of the object happen to exert pressure on the back of the eye at point 1, they in this way would pull the whole thread 1–2 and enlarge the opening of the tubule marked 2. And similarly, the rays that come from point B enlarge the opening of tubule 4, and so on with the others. Whence, just as the different ways in which these rays exert pressure on points 1, 3, and 5 trace a figure at the back of the eye corresponding to that of object ABC, so, evidently, the different ways in which the tubules 2, 4, 6, and the like are opened by filaments 1–2, 3–4, and 5–6 must trace [a corresponding figure] on the internal surface of the brain."

with a single anatomical structure. Thus Descartes' speculative physiology defined his visual anatomy. His achievement was in presenting an account of the visual pathways in terms of their topographical organisation.

Descartes' analysis of vision was based on his conception of light: when light strikes the eye it applies force to points on the retina which are transmitted along the optic nerve to the brain. The significance of Descartes' theories does not lie so much in whether they were valid or not as in the emphasis they placed on physiological interpretations of vision. Increasingly, from that time onwards texts that dealt with vision would have some diagram representing the pathways from the eyes to the brain, and some speculations regarding the site at which vision occurs. In the context of the visual pathways to the brain, it was believed that messages carried by the two optic nerves remained separate until they were united in the brain, as is indicated in Descartes' figure.

In Query XV of his *Opticks*, Newton asked whether the fibres from corresponding regions of each eye were united before they reached the brain, adding a telling reflection on species differences in the visual pathways to the brain. In fact, he had carried out experiments on optic nerves around 1682 but he did not publish details of them. Newton advanced our knowledge of the binocular pathways by describing the partial crossover of nerve fibres from one optic nerve to the other side of the brain. This discovery was used to support a mechanistic interpretation of vision, that is, one that did not make any appeal to non-material sources like the soul. The messages from each eye were thought to become one, because the nerve fibres themselves were (wrongly) considered to unite in the optic chiasm (see Figure 5.3). He also observed that there was almost complete

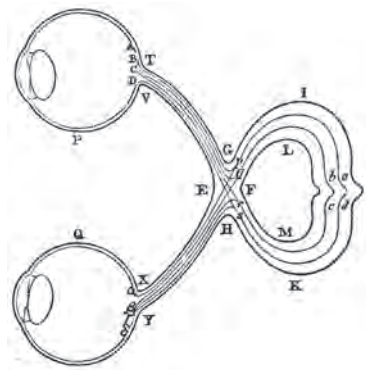


Figure 5.3 Binocular vision according to Newton. "Now I conceive that every point in the retina of one eye, hath its corresponding point in the other; from which two very slender pipes filled with the most limpid liquor, do without either interruption, or any unevenness or irregularities in their process, go along the optic nerves to the juncture EFGH, where they meet either betwixt G, F, or F, H, and there unite into one pipe as big as both of them; and so continue in one, passing either betwixt I, L or M, K, into the brain, where they are terminated perhaps at the next meeting of the nerves betwixt the cerebrum and cerebellum, in the same order that their extremities were situated in the retina's."

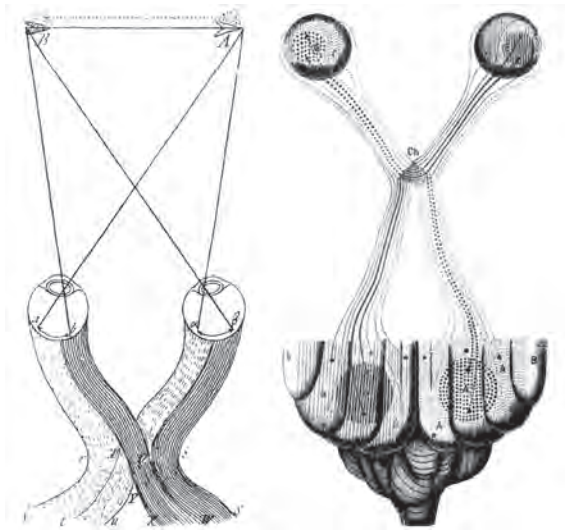


Figure 5.4 (a) Chevalier Taylor's diagram of the optic nerve and optic tract (from 1738). (b) Hermann Munk's diagram from 1879

crossing of the fibres for animals with lateral, as opposed to frontal, eyes. Newton was almost correct in his analysis: partial decussation was appropriate, but he represented the nerves themselves as uniting at the chiasm.

Optic nerve fibres from corresponding points on each eye formed single fibres in the optic tract. This detail was rectified by John Taylor (1708–1772) in an accurate representation of the partial crossing over and independence of the nerve fibres published in 1738 (Figure 5.4a). Taylor, who was a notorious oculist, probably based his diagram on shrewd speculation. Fibres in the optic nerve diverged after the optic chiasm, with those from the left halves of each retina projecting to the left part of the brain, and vice versa. This was not firmly established anatomically until the end of the nineteenth century by Hermann Munk (Figure 5.4b).

The comparison between the illustrations of Taylor and Munk is interesting as it indicates how much had been learned about the brain in the century and a half that separated them. Taylor's optic tract fibres go nowhere, unlike the detailed cortical projection represented by Munk. A modern schematic diagram of the visual pathways is shown in Figure 5.5.

The axons of the retinal ganglion cells leave the eye forming the optic nerve. This nerve is situated at the base of the brain, and the two optic nerves travel towards one another and appear to meet at the optic chiasm (see Figure 5.5); it was so called by Galen because it resembled the shape of the Greek letter χ (chi). The axons do not in fact meet, but they project to different cerebral hemispheres according to the area of the retina from which they originate: axons from the outer (temporal) side of each retina (the left half-retina for the left eye and the right half-retina for the right eye) project to the ipsilateral hemisphere (on the same

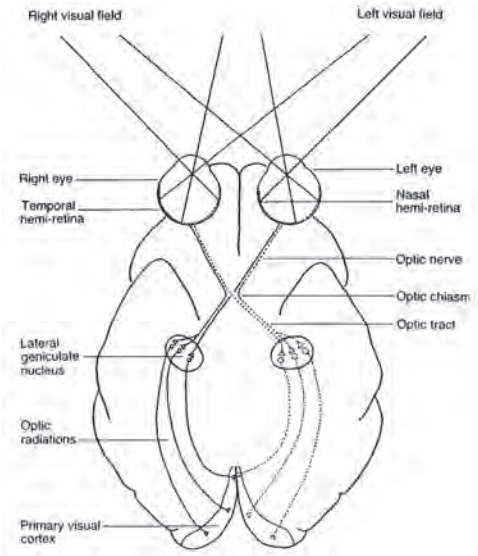


Figure 5.5 Pathways from the eyes to the visual cortex, viewed from below. The axons of the retinal ganglion cells form the optic nerves which pass under the frontal lobes of the brain. Fibres from the nasal halves of each retina cross over to the opposite hemisphere, while those from the temporal halves project to the hemisphere on the same side. This partial decussation occurs at the optic chiasm, and it results in signals from the same regions of the visual field projecting to the same hemispheres. The optic tracts consist of fibres from both eyes, and the axons terminate in a large subcortical body called the lateral geniculate nucleus (LGN). The cells from the LGN project to the primary visual cortex, also called V1.

side), and the axons from the nasal halves of the retinas cross over and project to the contralateral hemisphere. There is a narrow vertical strip in the centre of both retinas, subtending about 1 deg, that projects to both hemispheres. Animals that have lateral rather than frontal eyes, with little or no binocular overlap of their two visual fields, have almost complete crossover at the optic chiasm. The partial crossover of fibres in humans results in the transmission of signals from equivalent parts of the visual field to the same hemispheres. Thus, the right half of the visual field projects to the left halves of the retinas, and they in turn send neural signals to the left hemisphere.

The axons from the retinal ganglion cells continue beyond the optic chiasm as the optic tract, and they form synaptic connections in the thalamus, in a structure called the lateral geniculate nucleus (LGN). The LGN consists of six concentric layers, rather like an onion. The fibres from the contralateral eye project to layers 1, 4 and 6, and those from the ipsilateral eye to layers 2, 3 and 5 (see Figure 5.6). The cell bodies in layers 1 and 2 are larger than those in layers 3 to 6; the former are called magnocellular layers and the latter parvocellular layers. The midget (small) retinal ganglion cells project to the parvocellular layers, and the parasol

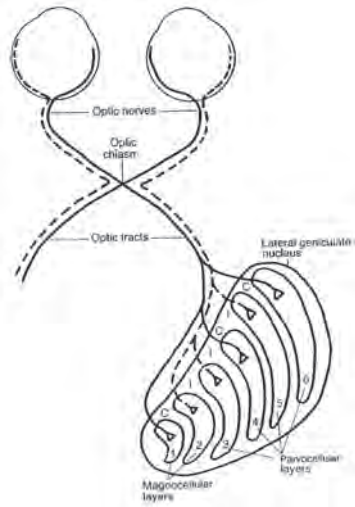


Figure 5.6 The lateral geniculate nucleus (LGN) consists of six layers. The cell bodies in layers 1 and 2 are larger than those in those in layers 3–6; the former are called magnocellular and the latter parvocellular layers. Projections from the contralateral (C) eye synapse in layers 1, 4 and 6, and those from the ipsilateral (I) eye synapse in layers 2, 3 and 5. Each layer can be further subdivided and very small koniocellular cells are located in the lower sublayers.

(large) cells to the magnocellular layers. Lying at the lower sublayers of the LGN are very small koniocellular cells; these receive projections from bistratified retinal ganglion cells.

Axons from the LGN cells project to the visual cortex, which is situated at the back of the brain. The visual cortex also has a vertical and a horizontal cellular organisation (Figure 5.7). Vertically (perpendicular to the cortical surface) it can be divided into six layers, on the basis of the cell types that can be distinguished microscopically. The horizontal organisation is in terms of the lateral fibre connections between ocular dominance columns.

Visual neurophysiology

Vision, like all other aspects of experience and behaviour, is mediated by activity in the brain. It also depends upon activity in highly specialised cells in the sense organs called receptors. A full understanding of vision will include an appreciation of the neurophysiological processes that are initiated by the activity of light on receptors in the eye. These involve the modification of light energy into nerve impulses and their transmission to the areas at the back of the brain where they are analysed. Tremendous advances have been made in our knowledge about the neural processes underlying vision, and many of these will be touched upon in the following sections.

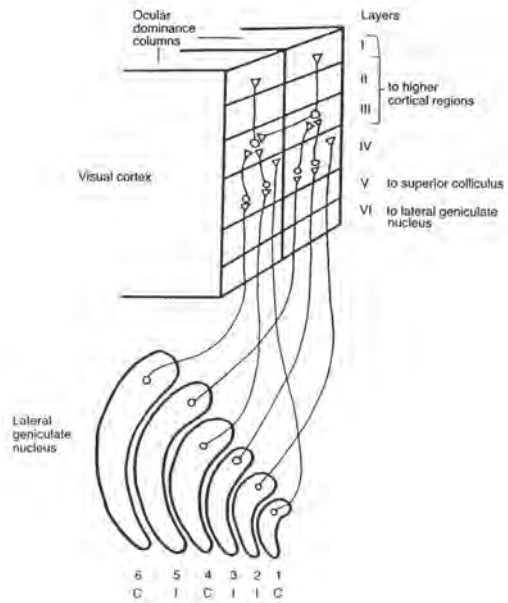


Figure 5.7 Projections from the LGN to the primary visual cortex (V1). All M and P cells project to layer IV of V1; those from the magnocellular layers project to higher locations in layer IV. Fibres from the ipsi- and contra-lateral layers of LGN project initially to adjacent regions of V1, maintaining a distinction based on the eye stimulated. There are also lateral connections between these adjacent regions. V1 has six distinct layers. K cells project to layers II and III Incoming fibres to layer IV project upwards to layers III, II and I, and are concerned with analysing the visual signals. Layers V and VI send signals back to subcortical structures, like to LGN from layer VI.

Psychology has constantly tried to link observable function (behaviour) with underlying structure (anatomy, physiology, genetics etc.), and these attempts are often considered to have been most successful in the area of perception. Some visual phenomena can be reduced to known neurophysiological processes. Reductionism, describing phenomena at one level in terms of concepts at a simpler level, has been one of the main motivations for scientific enquiry generally. It is certainly a powerful force in visual science. Paradoxically, we do not know a great deal about the neurophysiology of human vision at the cellular level, but we do know a lot about that in some of our nearest biological neighbours. Much of the material described below is based upon experiments on other species, particularly monkeys. We assume that equivalent processes occur in the human visual system because of the many biological similarities that exist: the receptors in the eyes, the pathways from the eyes to the brain, and the anatomy of the visual areas of the brain are all strikingly similar. Therefore, it does seem justified to relate the wealth of perceptual data on humans to the neurophysiological evidence from other species. Modern neural imaging techniques, like fMRI, are strengthening the evidence to support such

inferences. In this way it is possible to interpret visual phenomena in terms of the brain events that might give rise to them.

Neural activity in the retina

Having outlined the structure of the retina and the pathways to the brain, we can examine the activities of these cells. The initial stages of light absorption in the receptors results in hyperpolarisation (the inner part of the receptor becomes more negative with respect to the outer part). The measurement of such small graded electrical changes is possible by the use of minute pipettes (micropipettes) in contact with the surface of the cells or inserted into the cells themselves. The amount of hyperpolarisation that occurs is proportional to the light energy absorbed. The receptors form synaptic connections with bipolar cells and horizontal cells. The electrical activity of the bipolar cells depends on the source of stimulation: direct influence from the receptor cell results in hyperpolarisation, whereas indirect influence from horizontal cells produces depolarisation. The horizontal cells in synaptic contact with receptors also hyperpolarise. All these electrical events are graded, that is, they vary on a continuous scale according to the intensity of stimulation.

Quite a different electrical response can be recorded from the amacrine and retinal ganglion cells: they produce action potentials (or nerve impulses) which can be measured extracellularly with microelectrodes. Action potentials are all-or-none electrochemical events that have a fixed amplitude; they are the rapid depolarisations that occur along the axons of nerves. Much more is known about the activity of retinal ganglion cells (and subsequent stages in the visual pathway) because extracellular recording has been technically possible for longer than has intracellular recording. Indeed, the activity of retinal ganglion cells can be measured without entering the retina – by recording from single axons in the optic nerve or optic tract. This technique has generally been applied to experimental animals like cats and monkeys, and it has furnished us with important information concerning the ways in which patterns of light are coded in the visual system.

Prior to the use of microelectrode recording it was thought that the cells in the visual pathway were excited simply by the presence of light. Now it is known that it is the pattern of light that is of importance, not solely its presence. Experiments on *Limulus polyphemus* (the horseshoe crab) first indicated that the neural activity of one cell can be influenced by the state of its neighbours, a process called lateral inhibition. This is evident at the retinal ganglion cell level of vertebrate eyes. When experiments were performed to determine the adequate or appropriate stimulus for these cells it was found that they did not fire (produce nerve impulses) when the whole eye was stimulated by diffuse light. Firstly, light had to be presented to a particular part of the retina, and secondly it had to have specific dimensions. Some retinal ganglion cells could be excited by small spots of light, but they ceased firing when the spot was increased in size (Figure 5.8). These are called on-centre cells, as they are characterised by excitation when the centre is stimulated by light and by inhibition when the annular surround is also

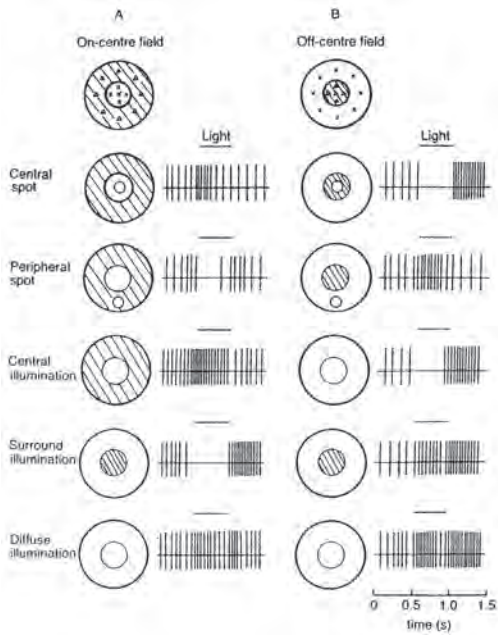


Figure 5.8 Responses of on- and off-centre retinal ganglion cells to a variety of light patterns falling on their receptive fields. The cross symbols in the upper figures refer to excitatory influences and the triangles signify inhibition.

stimulated. Retinal ganglion cells, like other neurons, have a resting discharge rate, i.e., they continue to produce nerve impulses at irregular intervals in the absence of light. Neural inhibition can be demonstrated by stimulating the surround alone with light, as the resting discharge rate then declines. Other retinal ganglion cells display the opposite pattern of activity: they are excited by the presence of light in the annular surround and inhibited by its presence in the centre. These are called off-centre cells.

The activity of each retinal ganglion cell can be influenced by a particular pattern of light falling on the appropriate part of the retina. This region is called the receptive field for that cell. Most retinal ganglion cells have concentric receptive fields with central and surrounding regions that are antagonistic in their function. Thus, the processes of neural excitation and inhibition are vital in determining the ways in which cells in the visual system respond to the patterns of light falling on the retina. The retinal ganglion cells respond to changes in the pattern of illumination, rather than to steady states of uniform illumination. These changes can be spatial and temporal. So far we have described the spatial characteristics – the centre-surround receptive fields, but there are differences in the responses of retinal ganglion cells to the wavelength of light, the duration of stimulation and the speed of conduction along their axons. Three main categories of cells leaving the retina have been identified and they are called midget, parasol

and bistratified retinal ganglion cells. Prolonged stimulation (for a few seconds) of the midget (small-bodied) retinal ganglion cells results in an initial reduction of the firing rate, levelling off to a steady rate, and they conduct slowly (at around 6 m/s). These have a sustained response; that is, they continue to discharge at above their resting rate with prolonged stimulation. Because they project to the parvocellular layers of the LGN they are considered to be part of the parvocellular (or P) pathway. Midget retinal ganglion cells make up about 80 per cent of the cells projecting to the LGN. The parasol (large-bodied) retinal ganglion cells respond briefly to the onset or offset of stimulation, and then revert rapidly to the resting discharge rate even if stimulation is continued; they also conduct nerve impulses more rapidly (at about 15 m/s). Thus, they have a transient response and project to the magnocellular layers of the LGN, and are part of the magnocellular (or M) pathway; they account for about 10 per cent of the retinal ganglion cells. As noted above, the P pathway receives input from predominantly in and around the fovea, where visual acuity is best, and they have smaller receptive fields than cells in the M pathway, which are distributed more evenly throughout the retina. The receptive fields of retinal ganglion cells are antagonistic with the centre stimulated by the opposite characteristics to the surround. Cells in the M pathway code luminance differences between centre and surround whereas those in the P pathway code wavelength as well as luminance differences. The K (koniocellular) cells, lying at the lower sublayers of the LGN, receive their input from the bistratified retinal ganglion cells, and they account for less than 10 per cent of the fibres leaving the retina. They respond to short wavelength (bluish) light by increasing their firing rate and to medium wavelengths (yellowish) by decreasing firing. Accordingly, they are considered to be involved in coding the lower wavelength light (blue and yellow), whereas the P cells respond to the long and medium wavelengths concerned with red–green coding.

The excitatory and inhibitory interconnections in the retina are the basis for the receptive field properties of the retinal ganglion cells. The antagonistic interaction between centre and surround also determines the type of stimulus that will produce a response in these cells. As we have seen, the maximum response is elicited by a small white or dark spot, but large responses will also be produced by the grating stimuli described in the previous section. That is, the light bar of a grating will produce a large response if its projected width on the retina corresponds to the dimensions of the on-centre receptive field; conversely, a dark bar of the grating would produce a large response in an off-centre retinal ganglion cell with a receptive field of the appropriate dimensions.

Neural activity in lateral geniculate nucleus

This distinction between large- and small-bodied retinal ganglion cells, in the M and P pathways, respectively, is amplified in the projection to the LGN (refer back to Figure 5.6): the small colour-opponent retinal ganglion cells project to the parvocellular region (layers 3–6) and the large cells project to the magnocellular region (layers 1 and 2). Therefore the large and small cells that were intermingled

in the optic tract project to separate sites in the LGN. Moreover, these separate M and P pathways display other functional differences in addition to wavelength sensitivity: P cells respond a) more slowly, b) to higher spatial frequencies and c) with less sensitivity to luminance contrast than M cells. In addition, there are K (koniocellular) cells which are distinct from the P and M cells. They have very small cell bodies and are not restricted to specific layers but are located at lower levels in them. It is believed that K cells are involved in colour processing, particularly concerning short wavelengths. The P system could be concerned with red–green processing and the K system with blue–yellow. In addition to sending signals to layer IV of the primary visual cortex, the LGN also receives input from layer VI (as is indicated in Figure 5.7).

Contrast phenomena

The lateral inhibitory interactions within the retina act to increase the differences between neural signals for edges or boundaries in the stimulus. Therefore, the neural coding of light falling on the retina depends on changes in its spatial distribution over time; these changes are enhanced by the neural machinery within the retina. Indeed, the enhancement is such that it can lead to the visibility of patterns that are not present physically in the stimulus. Mach bands are one phenomenon of this type; they are named after the physicist Ernst Mach (1838–1916) who described them in the nineteenth century (Figure 5.9). The areas of constant luminance on the left and right are separated by a ramp of increasing luminance: if a light meter was passed horizontally across the pattern it would register a constant level then a gradual increase followed by another constant (but higher) level of luminance. This does not correspond with what is seen in the

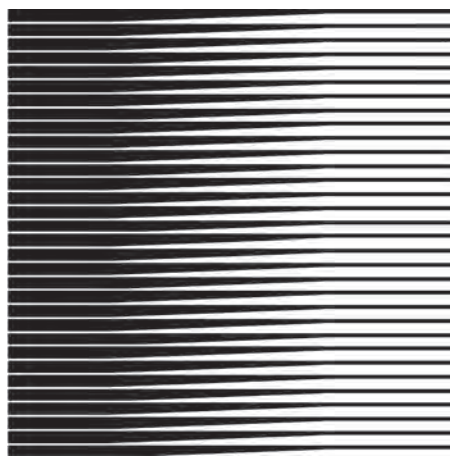


Figure 5.9 Mach bands refer to the light and dark bands that can be seen flanking the boundaries between the luminance ramp. A dark band is visible on the darker side of the ramp and a light band on the lighter side, despite the absence of such differences in the pattern when measured with a light meter.

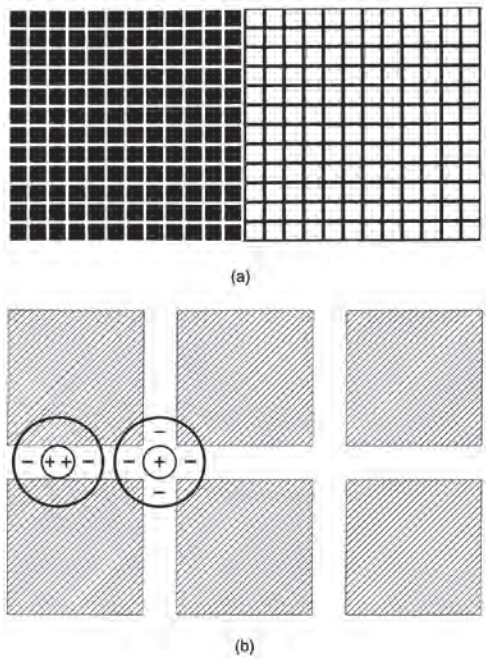


Figure 5.10 Hermann–Hering grids. (a) Dark grey dots appear at the white intersections on the left and light dots are visible at the black intersections on the right. Note that these might not be apparent at the intersection you fixate upon. (b) The interpretation of the Hermann grid effect in terms of on-centre receptive fields. The surround inhibition is greater at the intersections of the white lines than between the squares, because more of the surround is exposed to light. A similar interpretation of the Hering grid can be given in terms of off-centre cells.

pattern, as there appears to be a dark vertical band on the left side of the ramp and a light band on the right, which are called Mach bands.

Shortly after Mach had described his phenomenon (in 1865), a related effect was reported by Ludimar Hermann (1838–1914) when looking at a grid of black squares on a white background (Figure 5.10a): the background does not appear uniformly white, but dark grey dots are apparent at all the intersections apart from the one fixated. The reverse occurs with white squares on a black background, as was reported by Hering. These are now referred to as Hermann–Hering grids, and they have been related to the properties of concentric receptive fields in the retina. Consider two equivalent on-centre receptive fields stimulated by different parts of the white background in a Hermann grid (see Figure 5.10b). The one that falls between the intersections would be inhibited by a smaller white area surrounding it than will the one falling at an intersection.

The dimensions of concentric receptive fields are smallest in the retinal ganglion cells receiving their input from the central fovea, and they increase in size with increasing distance from the fovea. This is probably the reason why the

Figure 5.11 Simultaneous brightness contrast. The central circles are all the same luminance, but their brightnesses are influenced by the area surrounding them. If the surround is darker then the grey circle looks brighter, and vice versa.

Hermann–Hering dots are not visible at the fixated intersection: both the centre and surround would fall between the squares and within the intersections, so that there would be no differential response from them, and no illusory dots. If the separations are made sufficiently small then the dots can be seen at the fixated intersections. In fact, measuring the limiting dimensions of grids that yield the illusory dots has been used to estimate the sizes of receptive fields at different eccentricities in human vision.

Mach bands and Hermann–Hering grids can be encompassed within a wider class of phenomena, namely simultaneous contrast effects. These were described by Goethe and investigated in greater detail by Michel Eugène Chevreul (1786–1889); they generally refer to the apparent brightness or colour of one region in a pattern when it is surrounded by another having a different brightness or colour. For example, in Figure 5.11 the central grey circles all reflect the same amount of light, i.e., they have the same physical luminance, but there is a difference in their brightness. This difference in brightness (a perceptual dimension) where there is no difference in luminance (a physical dimension) is due to the surrounding regions in each case. When the surround is lighter than the central grey circle it appears darker than when the surround is darker than it. If the surrounds were coloured then the central grey areas would also appear slightly coloured, but in the complementary colour of the surround. For instance, if the surround was red then the physically grey centre would appear greenish; if it was blue then the centre would appear yellowish. These effects have often been manipulated systematically in art, particularly in the Pointilliste paintings of Georges Seurat (1859–1891).

Colour

Seurat also sought to capture colour on canvas as if it was light. Painters traditionally mixed pigments on a palette before applying them to the canvas; Seurat adopted another technique which is more akin to additive than to subtractive colour mixing. The colours were applied as small dots of (principally) primary colours; they could then mix in the eye like light. His influences were Chevreul and Helmholtz, both of whom had written on the relationship between painting and perception, and Helmholtz advocated a theory of colour vision based on three primaries. The Young–Helmholtz trichromatic theory of colour

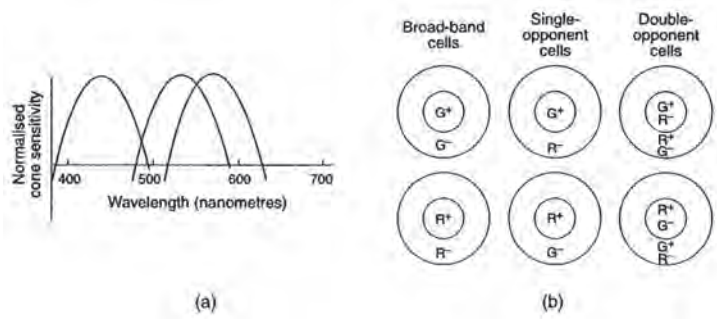


Figure 5.12 (a) Cone sensitivity functions for the three cone types, normalised to indicate the peak spectral sensitivities. (b) Different types of opponent process receptive fields found in the retina.

vision has received support from studies of the different photopigments in retinal cones. Both psychophysical and physiological studies have supported the view that there are three different types of cone receptors, with peak sensitivities at around 420, 530, and 565 nm (see Figure 5.12a). These are sometimes referred to as S (short), M (medium), and L (long) wavelength receptors and at other times, more loosely, as blue, green, and red receptors. Neither set of terms is strictly appropriate for the cone pigment that has a peak sensitivity at 565 nm, largely because it is neither a long wavelength in visual terms nor does it correspond to the perception of red.

The complement of colour cones varies markedly between species. The retinas of mammals, other than primates, have only two cone pigments, responding to short and medium–long wavelengths. Most Old World primates have three cone pigments, but some New World primates have four. It would seem that in humans the differentiation between the medium and long wavelength cone pigments is of relatively recent origin. The genes for the medium (green) and long (red) wavelength pigments are almost identical and they both lie on the X chromosome, which defines the sex of individuals. Females have XX chromosomes whereas males have XY, with the Y chromosome being shorter than the X. This can account for the wide divergence of colour defects in females and males, because a recessive on one X chromosome of a female is likely to be paired with a dominant allele; a recessive on the X chromosome of a male will have no allele with which to combine and so will be expressed (see Chapter 2). The short wavelength cone pigment is coded on a different chromosome.

The signals from the three cone receptors are combined neurally in many ways, as indicated by the receptive field properties of retinal ganglion cells in the P pathway. For example, some cells are excited when long wavelength (red) light falls on the centre and inhibited when the complementary (green) falls on the surround. All excitatory and inhibitory combinations of these colour opponent cells have been found with pairings between long and medium (red–green) and short and medium–long (blue–yellow) wavelengths (Figure 5.12b). Thus, there is

support for the Young–Helmholtz trichromatic theory at the receptor level, and for Hering's opponent process theory at the retinal ganglion cell level.

Neural activity in the visual cortex

The axons from the LGN project to visual area 1 (V1) of the visual cortex; the M and P fibres enter it in layer IV and the K fibres in layers II and III (refer back to Figure 5.7). Layer IV has been further subdivided anatomically and the distinctions between the P and M projections are preserved in the parts of layer IV in which they form synaptic connections. P fibres terminate in the lower layer (IVCβ) and the M fibres above them (IVCα). The cells form synaptic connections with neurons projecting vertically to upper layers of V1 (primary visual cortex). In addition to separate regions in layers I–III receiving input from P, M and K systems there is another which receives input from both. The regions are shown up by specific staining procedures as oval blobs separated from one another, and microelectrode recordings show that cells in these regions have quite different receptive field properties. In fact, almost all the cells in the visual cortex have receptive field characteristics that differ from the concentric organisation found at earlier stages in the visual system. The neurons in the blobs receive inputs from both the P and M systems: they can be excited by coloured or white light falling on the retina, but they are not as sensitive to its precise position on the retina. The cells between the blobs receive inputs from the P system: they are excited by edges or bars in a particular orientation, but are relatively insensitive to wavelength. That is, a cell will respond to a line at a specific orientation no matter what its wavelength or intensity, but it will not respond to lines that are inclined away from the preferred orientation (Figure 5.13). The cells in layer IV that receive their inputs from the M pathway are excited by lines or edges in a specific orientation, particularly if they are moving; they are not selective for wavelength. For example, a given cell might respond most strongly when a horizontal line moved vertically.

These orientation selective neurons were once considered to make up the majority of cells in the visual cortex. Indeed, such was the interest elicited after their discovery, by Hubel and Wiesel in 1959, that they became called feature detectors (see Chapter 3). This was because the cortical neurons were selectively tuned to extract certain features contained in the pattern of retinal stimulation – in this case orientation. Direction of edge motion is another feature extracted, as are binocular disparity and colour. The orientation selective neurons have a highly ordered organisation over V1, both perpendicular to the cortical surface and parallel to it, as is shown in Figure 5.14. If a microelectrode is inserted at right angles to the surface and all orientation cells encountered are recorded, then they will have the same preferred orientation, say 45 deg. This has been called a cortical column. An equivalent insertion in the neighbouring column will record a slightly different orientation selectivity, say 30 deg; the next column would have cells displaying a preferred orientation of around 15 deg, and so on, each column differing by about 15 deg from its immediate neighbour. In fact, the orientation

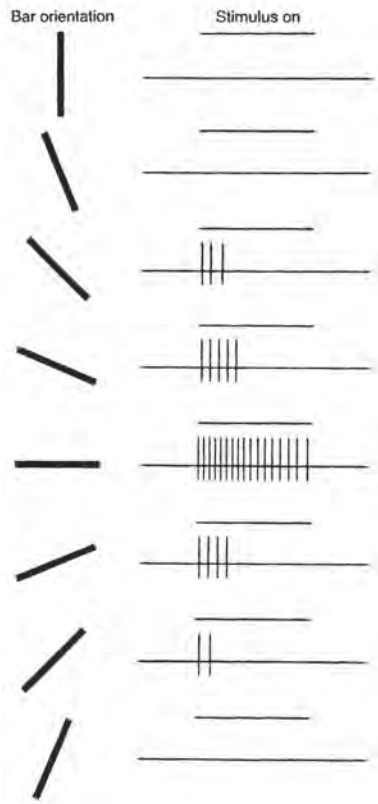


Figure 5.13 Receptive field properties of a cortical orientation detector. The orientation at which a bar is presented to a specific part of the retina is shown on the left. On the right the upper horizontal lines indicate when the stimulus (a black bar on a light background) is presented, and the vertical lines denote nerve impulses. For this cortical cell the firing rate is highest for a horizontal bar, and it declines sharply when the orientation differs from horizontal, not responding at all for bars at right angles to the preferred orientation. Neighbouring cells respond most strongly to different orientations, so that all orientations are represented for all retinal regions.

selectivity of the cortical cells is not quite as precise as might have been suggested here; there is a particular orientation that will produce the maximum response from each neuron, but it will also respond, though with decreasing intensity, to lines within 10–20 deg either side of the preferred orientation. Accordingly, a given line falling on the retina will excite a range of orientation selective neurons, but to varying extents. It is clear that area V1 has a vertical organisation in terms of orientation selectivity, and a horizontal organisation in terms of orientation change. This results in all orientations being represented for a given retinal region. The cortical blobs code the wavelength characteristics of the stimulus; the cells are non-orientation selective and they have double opponency. The receptive fields have a centre–surround organisation with the centre, say, being excited by

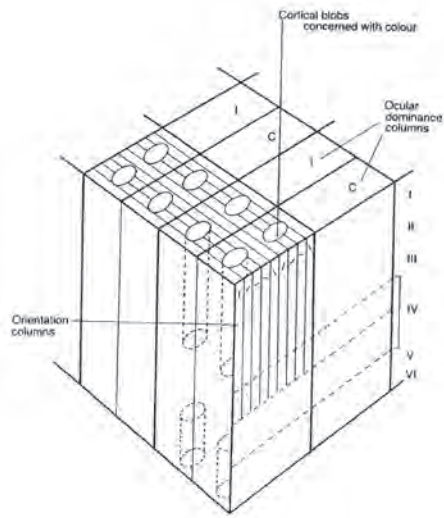


Figure 5.14 The organisation of cells in the primary visual cortex (V1). Cells in layers I, II, II and the upper part of layer IV respond to contours in specific orientations. A column, perpendicular to the surface of the cortex, contains cells having the same receptive field orientations, and is referred to as an orientation column. Neighbouring columns display a shift of preferred orientation of about 15 deg. There are also cortical blobs, receiving inputs from both magno and parvo cells in LGN, that respond to the wavelength of the original light stimulus. Projections from the ipsilateral (I) and contralateral (C) eyes remain separate in V1, resulting in a larger horizontal organisation based on ocular dominance. Within cortical areas of about 1 mm^2 both eyes and all orientations are represented for a small region of the visual field. Hubel and Wiesel referred to these as hypercolumns.

long wavelengths and inhibited by medium ones (denoted by R+G–) with the reverse influence on the surround (R–G+). These cell types are more common than those coding for B+Y– centre and B–Y+ surround or the converse.

So far, nothing has been said about combining signals from the two eyes. The fibres in the optic tract are from similar halves of both eyes, but they project to different layers in the LGN. The first stage at which binocular integration of the neural signals occurs is in V1. The cortical neurons can be excited by appropriate stimulation of either eye, although one eye will generally have a greater influence than the other. This aspect of eye preference represents yet another horizontal organisation over the surface of V1: neighbouring orientation columns tend to have the same eye preference. However, at intervals of about 0.5 mm the eye preference changes abruptly to the other eye, and the sequence of orientation selective columns is repeated, as is represented schematically in Figure 5.14. Figure 5.15 shows the pattern of eye preference regions that have been established in monkey visual cortex. The eye-preference slabs can be identified by particular staining procedures and they are striped, rather like the stripes on a zebra. All the cells in the regions shown in white have one eye preference and those shown in black have a preference for the other eye. The combination of binocular signals to

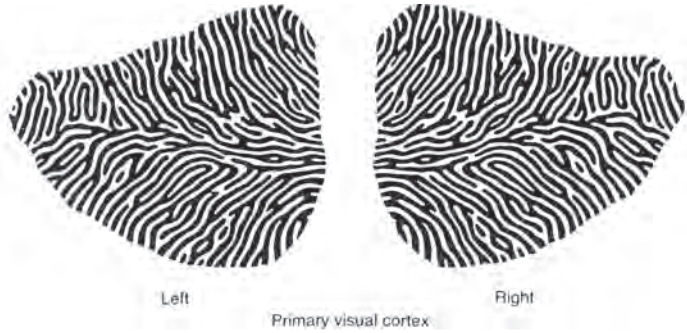


Figure 5.15 The pattern of ocular dominance columns, represented by black and white areas, over the left and right hemispheric regions of V1.

furnish information about relative depth (stereopsis) takes place at the next visual area.

V1 is, therefore, a highly organised structure. It appears to break down the pattern of light falling on the eye into discrete features, like retinal location, orientation, movement and wavelength, as well as maintaining a difference between the signals from the two eyes. These features are further differentiated in the subsequent cortical processing in visual areas shown in Figure 5.16c. Axons from neurons in V1 project to a variety of other visual areas in the cortex (Figure 5.16d). The adjacent visual area (V2) receives inputs from the three receptive field types in V1, but subsequent visual areas appear to specialise in the processing of particular features extracted in V1. Rather than having the blob configuration of V1, V2 is organised in stripes which are differentiated by their widths. Thick stripes receive their input from the M pathway (layer IVCα), thin stripes from the blobs in V1, and the interstripe cells from the K pathway (layers II and III of V1).

The binocular combination in V1 is primarily for inputs from corresponding areas of each eye. That is, most binocular V1 neurons are excited by cells from the two eyes with equivalent receptive field characteristics and equivalent retinal locations. In the thick stripes of V2 there are binocular cells that respond most strongly when the receptive field characteristics are slightly different. The inputs from each eye would have the same orientation selectivity, but the retinal locations associated with the receptive fields would be slightly different. These have been called disparity detectors because they appear to be responding to specific disparities (or horizontal retinal separations) between the stimulating edges in each eye. There are some disparity detectors at the level of V1. In addition to the analysis of retinal disparity further specialisations can be found in V2; thin stripes are mainly processing colour and the interstripes are more specialised for orientation. This pattern of refining the aspects of the stimulus that are processed continues to the other visual cortical areas. For example, area V4 is principally concerned with the analysis of colour, with cells in areas V5 (also referred to as

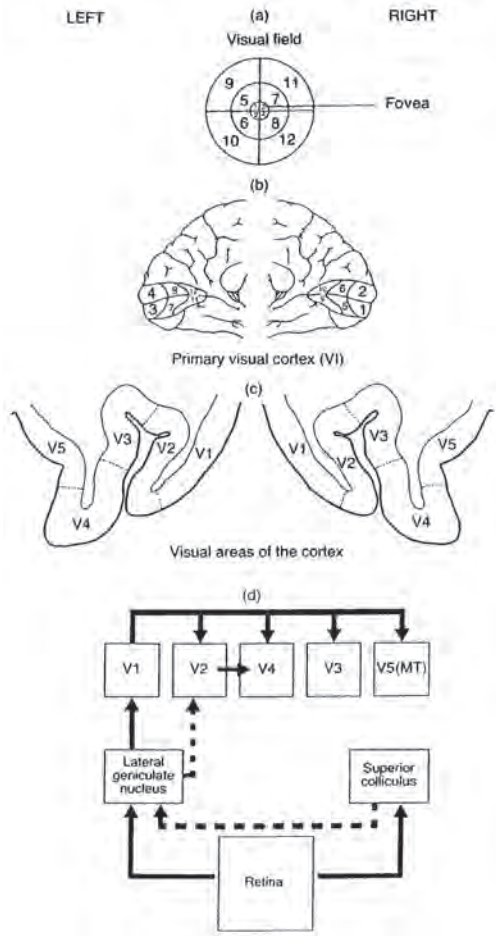


Figure 5.16 (a) The visual field considered schematically as 12 regions, centred on the fovea. (b) The cortical representations of these regions. The diagram represents the posterior portions of the occipital cortex spread apart, exposing area V1. Note the inverted and reversed relationship between the visual field and its cortical representation, and also the cortical magnification: the largest areas of V1 process signals from the small foveal area of the retina, corresponding to areas 1–4 in (a). (c) A horizontal section through the occipital cortex showing visual areas 1–5. Signals from LGN are processed in V1; V2 is principally involved in processing retinal disparities, V4 with wavelength and V5 with movement. (d) A schematic diagram of the connections between visual cortical areas.

MT) processing movement. There are even cells in a more central area that can process complex stimuli like faces.

The projections from the retina to the LGN and onwards to the multiple visual areas retain the basic mapping of the retinal surface: adjacent regions on the retina project to adjacent regions in LGN, V1, V2 etc. This is called retinotopic

projection because the distribution of stimulation on the retina is preserved at more central sites. In turn, the pattern of retinal stimulation is geometrically related to the layout of objects in space. While adjacent regions on the retina project to adjacent regions in the visual cortex, the precise separations are not retained. The fovea is only a small area in the retina, but its projection to the cortex is disproportionately large. Conversely, the large peripheral areas of the retina are represented by small areas in the cortex (see Figure 5.16a and b). If you think of the retinal surface as a balloon with a very weak central area (analogous to the fovea) then inflating it results in a huge expansion of the central region, but relatively little of the surrounding area. Neighbouring regions are still adjacent, as in retinotopic projection, but the relative separations are changed. This is a general scheme found in the cortical projections of sensory systems: those parts of the sensory surface that have the greatest sensitivity also have the greatest cortical projection. For example, the cortical representation of the sensitive fingers is far more extensive than that of the arms, even though the area of the skin surface is much less. Similarly with vision, the retinal region serving the sharpest acuity (in and around the fovea) also has the largest cortical representation. It has been estimated that the central 10 deg of the retinal surface is represented by about 80 per cent of the cells in the visual cortex. This concentration of cortical processing to the central visual field would present a problem if the eyes remained stationary, because stimuli of significance could move out of the region where acuity is highest. In fact, the problem does not arise because of our exquisite control over eye movements, so that we can pursue moving objects to keep them projecting to the fovea or we can move our eyes rapidly and accurately to locate novel objects in the visual field. These control processes are mediated by pathways in the midbrain that have not yet been mentioned.

Midbrain structures associated with vision

The discovery of multiple mapping in visual areas of the cortex is relatively recent, and as many as thirty such areas have now been found in monkey cortex. However, the concept of a second visual system was first proposed in the 1960s. Unlike the multiple maps in the cortex, this one is subcortical and older in evolutionary terms. The pathways that have been described so far comprise the classical projection system. Some fibres branch off the optic tract before the LGN to make synaptic connections in a structure in the midbrain called the superior colliculus (Figure 5.17). The paired superior colliculi also receive inputs from layer V of V1. In evolutionary terms the superior colliculus is far older than the visual cortex. In many animals, like reptiles, all visual processing occurs in the optic tectum (which is equivalent to the superior colliculus in mammals), as there is no cortex present. In mammals its function is related to localising objects in peripheral vision and directing eye movements towards them. Experimental animals that have had the visual cortex removed can still locate objects on the basis of collicular activity, but they do not demonstrate recognition behaviour when tested.

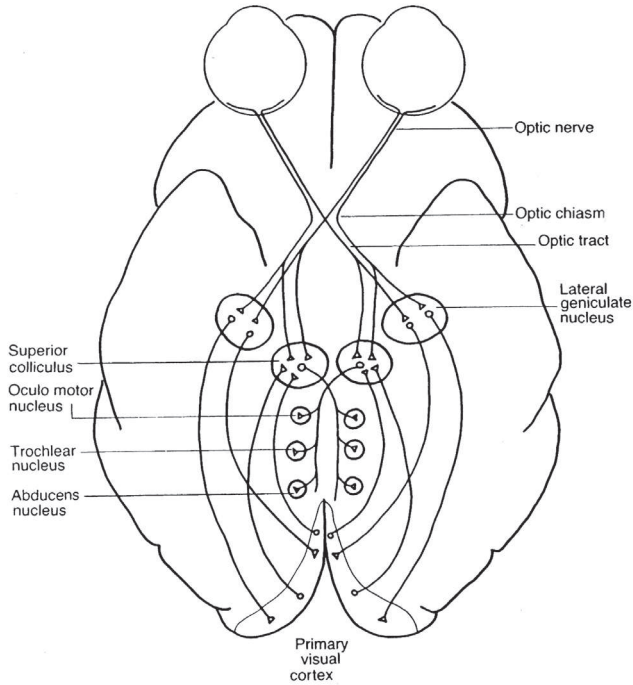


Figure 5.17 The classical projection pathway from the eyes to the brain via LGN, and the older pathway, in evolutionary terms, to the superior colliculus. Fibres from the optic tract synapse in superior colliculus, and the collicular fibres project to three nuclei concerned with the control of eye movements. Note that the superior colliculus also receives feedback from (layer V of) V1.

The midbrain also contains a number of oculomotor nuclei that control the movements of the extraocular muscles within the orbit. All eye movements are rotations generated by the coordinated activities of pairs of muscles attached to the eye ball. There are three pairs of muscles which can rotate the eye in the three dimensions of space (Figure 5.18). We commonly refer to eye movements in terms of the directions the eyes point rather than the axis around which they rotate. The most frequent movements are up and down, and right and left. The former are called conjugate eye movements because they are coordinated in the two eyes, i.e., both move up and down in the same way. Movements to the left and right can be conjugate, as when following a laterally moving object; these are called version eye movements. When the eyes follow an approaching object the movements are non-conjugate: one eye moves to the left and the other to the right and these are called vergence eye movements. As an object gets closer the eyes converge whereas when it recedes they diverge.

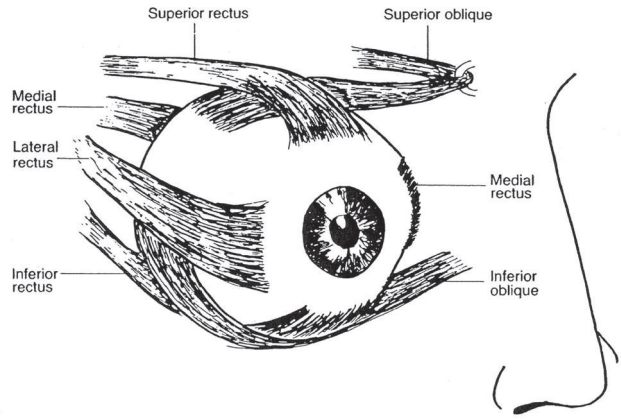


Figure 5.18 Extraocular muscles of the right eye. Rotations of the eye about a vertical axis (left and right) are controlled by the medial and lateral rectus muscles, those about a horizontal axis (up and down) by the superior and inferior rectus muscles, and those about the optical axis (torsion) by the superior and inferior oblique muscles

Neurophysiological interpretations of visual phenomena

Very little is known about the neurophysiology of the human visual system at the cellular level. Most of the experimental findings summarised above have been derived from experiments on many animals, but principally cats and monkeys. It might be the case that the human visual system behaves in quite a different way. This is, however, very unlikely. The anatomy of the visual system in humans is very similar to that of monkeys, and some species (e.g., macaque monkey) have about the same visual acuity and colour discrimination as humans. Therefore, we assume that the physiological processes measured in related species also apply to our visual systems. This extrapolation seems reasonable because the processes in cats and monkeys are closely comparable, even though the differences between their brain structures are greater than those between humans and monkeys.

Some phenomena have already been described in terms of the physiological processes that might give rise to them. For instance, the phases of dark adaptation were attributed to differences in rod and cone recovery times, or Hermann–Hering grid effects were related to differences in surround inhibition operating on the receptive fields of retinal ganglion cells. In attributing the phenomena to these particular processes we should not think that the phenomena are experienced at the levels these processes occur; the phenomena require those retinal processes but it is likely that they can only be experienced if the information is transmitted to more central sites. In order to sustain neurophysiological interpretations we need to assume that changes coded at one level in the visual pathway are retained throughout subsequent processing. The search for neurophysiological interpretations of visual phenomena has proved to be a driving force in vision research over the last few decades, especially since the discovery of feature

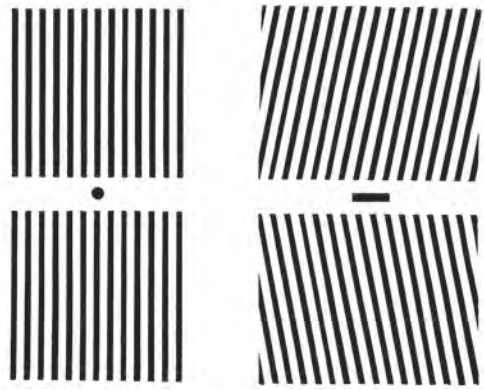


Figure 5.19 Stimuli for producing a tilt aftereffect. Initially look at the black spot on the left: the gratings above and below it will appear to be vertical and aligned. Then inspect the figure on the right by looking at the black rectangle for about 1 min; move your eyes slowly back and forth along the rectangle to avoid generating an afterimage. Looking back at the spot on the left will result in the lines appearing slightly tilted with respect to one another: the upper grating will look tilted anticlockwise, and the lower one clockwise. This aftereffect will only last for a few seconds after which the gratings will again look vertical and aligned

detectors in the visual cortex. The phenomena interpreted in this way tend to be rather simple, and they reflect the operation of the relevant stimulus dimension (like orientation or motion) in isolation.

Under normal circumstances we can judge orientation very accurately. Even in a darkened room we can adjust a light line to within about 1 deg of the direction of gravity (vertical). This accuracy in perceiving the vertical can be biased by observation of a non-vertical pattern. For example, the gratings on the left in Figure 5.19 will initially appear to be vertical and aligned. Following inspection of the pattern on the right for about 1 min, the gratings on the left will appear neither vertical nor aligned. This is called the tilt aftereffect, and it has been interpreted in terms of orientation selective cortical neurons. According to this interpretation, a vertical grating will stimulate a range of orientation selective cells in the visual cortex – those with preferred orientations inclined slightly away from the vertical as well as those for vertical (see Figure 5.20). Initially the balance for those inclined clockwise of the vertical will match that for those inclined anticlockwise. During inspection of the inclined grating an overlapping distribution of orientation selective cells will be stimulated. However, since the inspection is prolonged these will undergo adaptation, i.e., their firing rates will be reduced by constant stimulation, and the cells will require some time to return to their original sensitivity. When the original pattern is observed again, the balance between the clockwise and anticlockwise influences will be disturbed and shifted in the opposite direction to the inspection orientation. This shift in the distribution has been used to account for the shift in orientation that we can see. The effects of adaptation for seconds or minutes are relatively short lived, and so the two

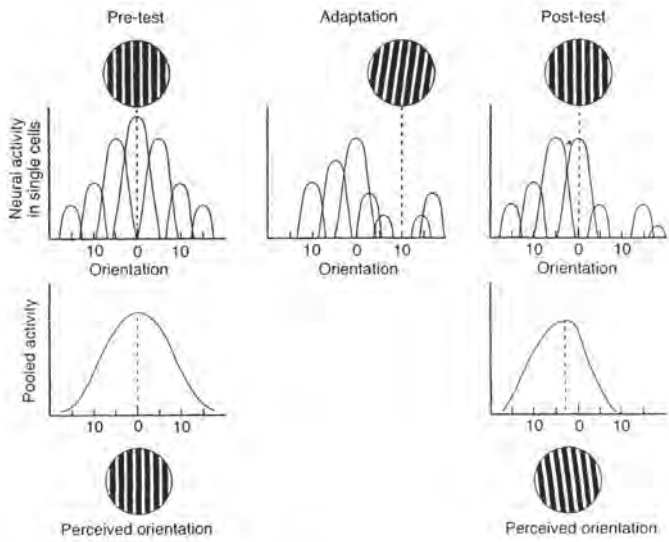


Figure 5.20 An interpretation of the tilt aftereffect based on adaptation of cortical orientation-selective cells. The vertical pre-test grating has the maximum neural effect on those cortical cells with vertically oriented receptive fields, but there will be smaller effects on other cortical cells having receptive fields within about 15 deg clockwise and anticlockwise of the vertical. The pooled activity of these cells is shown on the lower left. The effects are symmetrically distributed around the vertical, and it is assumed that the peak of this distribution represents the perceived orientation of the grating. During adaptation to a grating inclined 10 deg clockwise of the vertical, the cortical responses will adapt as shown in the central figure. On post-test, the vertical grating will stimulate some cells (clockwise of the vertical) which have been adapted and others (anticlockwise of the vertical) that have not. The resulting pooled distribution of activity (shown on the lower right) has its peak biased in an anticlockwise direction, with a consequent shift in perceived orientation. The adapted cells recover quite rapidly to their normal sensitivity and so the aftereffect also dissipates.

gratings return to alignment after a few seconds; longer adaptation results in longer lasting aftereffects.

The tilt aftereffect represents an example of successive biasing: some orientation selective cells are adapted and the recovery from this adaptation is evident in the misjudgement of orientation. Related orientation biases can be induced simultaneously in patterns like that in Figure 5.21. The grating in the cental regions is vertical , but it does not appear so becaused of the inclined grated surrounding: the vertical lines appear to be tilted in the opposite direction to their surrounding contours. This tilt illusion has been interpreted in terms of inhibition between orientation selective cells in the cortex.

The motion aftereffect is an example of motion biasing: following observation of a moving pattern (like descending water) a subsequently viewed stationary pattern appears to move in the opposite direction (see Chapters 2 and 7). It has been interpreted in a similar way to the tilt aftereffect, but incorporating motion

Figure 5.21 Tilt illusion. The central lines are parallel to the edges of the page, but they look tilted clockwise because of the surrounding anticlockwise lines.

detecting cortical cells rather than orientation detecting cells. Studies of tilt and motion aftereffects have also provided evidence for the ways in which signals from the two eyes interact. If a moving pattern is observed by one eye and a stationary pattern is subsequently observed by the other then the movement aftereffect can still be measured, but it only lasts for about 50 per cent as long as adapting and testing the same eye. This is called the interocular transfer of the aftereffect, and it has been used to infer the nature of binocular interaction in the human visual system. For example, the fact that interocular transfer occurs at all suggests that there are binocular cells that share the same information: they can be stimulated equivalently by either eye. The fact that transfer is less than complete suggests that there are also some monocular cells contributing to the perception. In addition to the binocular cells that are excited if either eye is stimulated there are others that are only excited if both eyes are stimulated at the same time. Some of these respond only when contours fall on corresponding locations in each eye, that is, when the receptive fields in each eye are exactly in register. Such binocular cells probably mediate binocular single vision – the perceptual experience of a single world. Other cortical cells respond only when there are small horizontal displacements between the locations of stimulation on each retina. These are called disparity detectors because they may be implicated in stereoscopic depth perception or stereopsis. Binocular single vision and stereopsis will be discussed in the next chapter. Interest in stereopsis has increased in recent years because neurophysiologists have found regions in the brain, like V2, which appear to extract and process retinal disparities in comparative isolation. Other brain regions process features like colour and movement. However, we know relatively little about the ways in which these ‘feature maps’ are subsequently integrated to provide us with unified percepts of objects having a particular shape, size, texture and location in space.

Neurophysiological investigations are conducted under conditions that are far more controlled than those operating in perceptual experiments. The animals

are generally immobilised so that the experimenters can define precisely both the area of the retina stimulated by light and the response that is elicited in the brain. Perceptual experiments, involving active and intact subjects, are open to much greater variability. The eyes move around resulting in the stimulation of a range of retinal regions, and the responses, like pressing a switch, are much more complex. There are several strategies that can be adopted to accommodate these differences. One is to try to simplify the stimulus and response conditions in perceptual experiments to match, as closely as is feasible, those obtaining in neurophysiological experiments. This is the approach that has been adopted in the study of spatial aftereffects described above. It has the advantage of proposing and testing explicit hypotheses linking neurophysiological and psychophysical data. The disadvantage is that the phenomena investigated are very simple, and the conditions of their occurrence are far removed from perception of objects in the world. Another strategy, and the one adopted in this book, is to examine functional aspects of perception, like location, motion and recognition, and to draw on the neurophysiological evidence when it is pertinent. This will enable the investigation of both simple and complex phenomena, but with respect to the latter it is more difficult to draw on our knowledge of neurophysiological processes underlying perception.

Visual processing beyond V1

As indicated in Chapter 1, a range of techniques is now applied to examining visual processing in addition to single unit recording, like EEG, ERPs, MEG, TMS and fMRI. These collectively have disclosed patterns of cortical processing that extend far beyond primary visual cortex (V1). Figure 5.16d indicates schematically the neural communication between the anatomical areas most closely associated with V1. Together these are called the striate (striped) cortex; the further areas which receive projections from it are called extrastriate regions, and there are many of them. It is as though the pattern of light falling on the retina is mapped many times throughout the visual system, with areas specialised for different dimensions of the pattern (like orientation, colour, stereoscopic depth, and motion). Indeed, the number of cortical maps concerned with visual processing continues to increase.

Evidence regarding the functions of extrastriate areas derives from two general approaches. On the one hand, techniques like fMRI enable determination of regions in which neural activity correlates with particular phenomena. For example, if a certain area is more active when the motion aftereffect is experienced then that area is implicated in motion perception. On the other hand, studies of the problems encountered by people who have had specific lesions to parts of the brain suggest the functions that those damaged parts serve. Both approaches suggest that there are two broad functional distinctions (visual streams) that start from the visual cortical areas. One passes dorsally to the posterior parietal areas of the cortex and the other ventrally to the inferotemporal cortex. These streams have been called the dorsal and ventral streams, and they are also referred to as the parietal and temporal streams; they could be involved in parallel processing

of the visual input. Lesions or sections in the posterior parietal cortex (involving the dorsal stream) result in failure to discriminate the spatial location of objects whereas lesions in the inferotemporal cortex (damaging processing in the ventral stream) impair object recognition. Accordingly, perceiving the location of an object was attributed to the dorsal stream while the ventral stream determined its identity – the *where* and *what* dimensions of vision. One version of the functions served by these streams is in terms of action and perception: the dorsal stream is said to be concerned with motor control and the ventral stream with perceptual representation. It has been argued that the representation and conscious awareness of external objects is a consequence of processing in the ventral stream. That is, our experience of a perceptual world is a consequence of this neural pathway. The control of our actions, on the other hand, is said to be due to the dorsal stream and it operates without our conscious awareness. The distinction between the two streams is becoming increasingly blurred, and their independent operation is constantly questioned.

Visual agnosias

A source of evidence for the notion of visual streams derived from the phenomenon of blindsight. It has this enigmatic label because the patients so afflicted do not seem to be consciously aware of the objects they respond to. When parts of the primary visual cortex have been damaged or surgically removed, patients do not detect stimuli presented to areas of the visual field that would project to the damaged cortex; the area of blindness is called a scotoma. They can see normally in the other parts of the visual field, but are unaware of light in the blind region. If asked to orientate themselves towards a stimulus in the scotoma they say they are guessing, but their responses are much better than chance. Thus, despite their blindness they can point to 'unseen' objects, indicate the direction in which they have moved, determine their intensity and even aspects of colour. The ability to orient towards objects that cannot be consciously reported suggests that information about objects in space is processed in more than one pathway from the eyes to the brain. The phenomenon indicates that information that is not consciously available can still elicit behaviour. This is one of the reasons why action and perception have been dissociated from one another.

Brain damage can have devastating effects on behaviour. Most of us have seen the consequences of a stroke on people, when one side of the body is partially paralysed due to injury of the brain on the opposite (contralateral) side. The restriction of oxygen to part of the brain results in death of some brain cells, and because of the separate blood supplies to the left and right hemispheres the brain damage tends to be restricted to one side. When the damage is to the left hemisphere of right-handed people there is generally some language disorder, whereas damage to the right hemisphere disturbs spatial orientation and recognition. For left-handers, this hemispheric specialisation is about equally left and right. Such observations of gross differences in the effects of damage to the left and right hemispheres led to the concept of cerebral lateralisation,

with the left hemisphere generally referred to as dominant because of its link to language.

Brain injury frequently has wide-ranging effects on behaviour because the damage is itself often widespread. However, some injuries to specific areas of the brain can manifest themselves in particular perceptual disorders, and important insights into the neural bases of behaviour can be derived from them. The cases are usually very rare but their implications are so profound that they demand detailed attention. Occasionally, the perceptual disorders can be most unusual, even bizarre, as in the tragic case of Dr P, graphically described by Oliver Sacks in his book *The man who mistook his wife for a hat*. This was an instance of prosopagnosia (the inability to recognise faces) and the title indicates that it applies to familiar faces; it can even be difficult for individuals so afflicted to recognise mirror images of themselves. Prosopagnosia is a special case of visual agnosia, which refers to the inability to recognise objects. For these conditions to operate, the failure of object recognition should not be due to damage to the sense organs or to any intellectual impairment. There are other specific visual agnosias that can affect recognition in different ways; one person was able to recognise the parts of objects but not the whole object itself, another was said to recognise living as opposed to non-living things, yet another could not see moving objects. However, the deficits are rarely as specific as these descriptions suggest, and the ways of classifying visual agnosias themselves are still contentious. The conditions are so rare that much of the evidence often rests on a small number of cases. Blindsight, for example, is critically dependent upon how specific the lesions are to V1, and so there should always be an element of caution in extrapolating from the investigation of single cases.

One of the distinctions made earlier – between action and recognition – is relevant to visual agnosias. When the injury is in the upper central part of the brain (the posterior parietal cortex) there are difficulties with reaching for and grasping objects, even though they can be recognised. With injury to the central and lateral region of the brain (the inferotemporal cortex) grasping is normal but recognising objects is disrupted. It is as though action is affected in one case and recognition in the other. One obvious issue raised by these patients is: what do we mean by recognition? Can a patient reach to an object that is not recognised at some level? The deficits in action and recognition associated with brain damage in humans have been supported by experimental studies with primates, where the location of lesions can be specified more accurately and thus provide invaluable clinical knowledge for recovery from such injuries.

Neuropsychological studies of individuals who have suffered accidental injury to specific regions of the brain have yielded fruitful hypotheses concerning the functions served by cortical areas and streams. The pace of change in the neuroscience of vision is so rapid that the current hypotheses are bound to be modified or replaced. However, as we will argue in the following chapters, both action and recognition systems can only operate effectively if there is a representation of three-dimensional space available to them. We will suggest how such a representation might be achieved.

Reference notes

Books on visual anatomy and physiology mentioned in the reference notes to Chapter 4 provide good illustrations of the brain and the pathways from the eyes to the visual cortex. The second edition of Milner and Goodale's (2006) *The visual brain in action* provides a general overview of the visual system, in addition to advancing a particular theory about its function (mentioned at the end of this chapter). Early views of the visual pathways can be found in Crone (1992), Polyak (1957) and Wade (1998a). Galen's book *On the usefulness of the parts of the body* has been translated into English by May (1968) and Descartes' *Treatise of man* by Hall (1972). Russell (2010) surveys the transition of ideas about seeing from Galen to Ibn al-Haytham. Newton's diagram of the visual pathways was neither published by him nor in the form shown; it was copied from his notes and Figure 5.2 is taken from Brewster (1855); Newton's text was published by Harris (1775). Chevalier Taylor not only provided the diagram of partial decussation, but his exciting life has been the source of many anecdotes (see Wade, 2008b). Many of the subsequent illustrations in this chapter have been redrawn from Kandel, Schwartz and Jessell (2000) in order to shown the basic structures within the visual brain as simply as possible. The book not only provides a thorough account of visual neuroanatomy and physiology but also relates structures to the possible functions they might serve.

Contrast phenomena have been a feature of pictorial art for much longer than their investigation in visual science would suggest (see Kemp, 1990; Wade, 2007a). Mach's studies of border contrast have been translated by Ratliff (1965), who drew attention to the use of the phenomenon in art. The novelty of Mach's figures were the ramp effects (like Figure 5.9) and similar brightness enhancement generated by rotating patterns. By contrast, the Hermann grid had its origins in sound rather than vision (see Wade, 2005b). Spillmann (1994) has argued that dimensions of human 'perceptive fields' can be derived from experiments on Hermann grids. Chevreul (1839) studied simultaneous contrast in the context of dyeing fabrics before it was investigated in the visual laboratory (see Kemp, 1990). Chevreul also produced an example of border enhancement that is often attributed to Mach: when rectangles of differing greys are juxtaposed, bands are seen at either side of the border (see Geier and Hudák, 2011). Many remarkable examples of simultaneous contrast can be found in modern books, as in Gilchrist (2006) and Purves and Lotto (2010). In his *Popular lectures on scientific subjects*, Helmholtz (1873, 1881) discussed not only the eye as an optical instrument but also the relation of optics to painting. In the latter he discussed colour and this essay is said to have influenced Seurat. The use of colour by artists and the changes that can take place in their colour vision with age have been analysed by Kemp (1990), Livingstone (2002), Werner (1998) and Werner, Pinna and Spillmann (2007). Mollon, Pokorny and Knoblauch (2003) provide a comprehensive survey of the mechanisms involved in normal and defective colour vision. Mollon has worked extensively in primate colour vision and the evolution of colour processing in mammals is described by

Jacobs (2009). Midbrain structures involved in vision are colourfully described in Morgan (2006).

Visual aftereffects have been referred to as the psychologist's microelectrode (Mollon, 1974; Frisby and Stone, 2010) because they are used to link quantifiable visual phenomena with the putative underlying neurophysiology. The tilt aftereffect is one of many that have been used in this context, and the interpretation illustrated is based on Blakemore (1973). As the name indicates, tilt aftereffects are successive shifts in visual orientation whereas the simultaneous biases are called tilt illusions, which have been subjected to similar interpretations (Daini, Wenderoth and Smith, 2003; Westheimer, 2011). A wide range of other orientation illusions are presented in Vicario (2011).

The neurophysiology of vision is undergoing constant change, with new techniques and discoveries. The excitement of research in this area of visual physiology is conveyed in the well-illustrated book by Hubel and Wiesel (2005). Frisby and Stone (2010) and Mather (2008) provide detailed and well-illustrated accounts of visual neurophysiology, relating the various cortical structures to aspects of vision. They also describe visual maps as do Wandell, Brewer, and Dougherty (2005) and Bridge (2011). Ungerleider and Mishkin (1982) initially proposed the distinction between visual streams signalling where and what. This division was modified to one of action and perception by Goodale and Milner (1992, 2005; Milner and Goodale, 1995, 2006). The debate concerning the validity of the latter model is still ongoing (see Westwood and Goodale, 2011; Schenk, Franz and Bruno, 2011).

Blindsight in monkeys and humans has been studied extensively by Weiskrantz (1986, 2009) and Humphreys and Riddoch (1986) have provided detail accounts of vision in a person with a specific visual agnosia. The popular books by Oliver Sacks (1985, 1998, 2009) have awakened broader interest in a range of disorders. Many texts are dedicated to clarifying the links between brain function and disorders of vision. Those by Farah (2004) and Goodale and Milner (2005) are particularly informative and present their respective theoretical interpretations of blindsight, visual agnosias and other perplexing departures of vision from the normal. Disorders of face perception as a consequence of brain damage are fully discussed in Young (2011) and they are integrated with fMRI studies by Kanwisher and Barton (2011). The specialised neural cells that respond to faces are described by Rolls (2011).

6 Spatial location

Our perceptual-motor coordination relies upon localising objects accurately, so that we can step over them, reach for them, avoid them or orient appropriately to them. Perception provides a platform for action, and actions take place in a three-dimensional environment. It is necessary, therefore, for spatial perception to represent the three-dimensional space in which behaviour occurs. The example of crossing the road, described in the first chapter, illustrated some of the ways in which the perception of location is necessary before other aspects of a scene, like motion or recognition, can be determined. This chapter will examine the perception of location in three-dimensional space, which requires information for the direction and orientation of objects as well as their distance away from us. These will be considered separately, although they almost always function together: information that an object is a given distance away is of limited use for guiding behaviour if its direction is not also detected, and vice versa.

Any statement about the locations of objects requires the specification of a frame of reference relative to which those locations can be assigned. This may seem self-evident, but investigators have often ignored the frames of reference operating in perceptual experiments. The empiricist philosopher, John Locke (see Figure 3.2), gave a very clear and instructive analysis of how frames of reference determine how we can speak about an object's location, independently of the object being judged. He imagined a chessboard with the pieces in play, located in a cabin of a ship at sea:

> Thus a Company of Chessmen, standing on the same squares of the Chessboard, where we left them, we say are all in the *same Place*, or unmoved; though, perhaps, the Chessboard hath been in the mean time carried out of one Room into another, because we compared them only to the Parts of the Chessboard, which keep the same distance one with another. The Chessboard, we also say, is in the *same Place* it was, if it remains in the same part of the Cabin, though, perhaps, the Ship it is in, sails all the while: and the Ship is said to be in the *same Place*, supposing it kept the same distance with the Parts of the neighbouring Land; though, perhaps, the Earth hath turned round; and so both Chessman, and Board, and Ship, have every one *changed Place* in respect of remoter Bodies.

The fundamental point that Locke was making is that statements about position, or what he called place, require the definition of a frame of reference, and that there are many such referents, which can be nested within one another. First it was the chess piece, then the board, then the cabin, then the ship, then the earth, then the solar system. Any statement about physical location is meaningless unless the frame or frames of reference are specified. Similarly, any perception of location necessarily requires a perceived frame of reference for that perception. This is equally true when considering the perception of changing location, i.e. motion, which will be the topic of the next chapter.

Accordingly, when discussing any aspect of space or spatial perception it is necessary to define the frame of reference relative to which measures are made. This is commonplace for physicists, who are accustomed to defining the framework within which any measurement is expressed. Surprisingly, it is less common among perceptual psychologists, even though they are dealing with a system that has multiple articulations. Light may stimulate receptors at a particular location on the retina of an eye, but that eye can rotate in the orbit, thus potentially changing the location of stimulation. What is true of one eye is true of both; in addition, the two eyes can move with respect to one another, and they can be displaced by head and body movements. Information about where the image of an object falls on the retina could therefore alter because the object moves, because the eyes move in the head, because the head moves with respect to the body or because the body moves in space. Any of these changes, alone or in combination, will result in the displacement of a retinal image. Despite this, our perception of objects is generally not disturbed and we perceive a stable world while complex patterns of eye, head and body motion take place. This is a very remarkable achievement of our perceptual system and it is one which is shared with other organisms that have mobile sense organs. A great deal of research on visual perception involves a stationary observer gazing at a fixed point, which eliminates the mobility that characterises normal perception and thus obscures its crucial role.

A key historical debate in cosmology illustrates the importance of distinguishing between physical and perceived frames of reference. Ptolemy proposed a geocentric model of the earth and its relation to the heavenly bodies. It was called geocentric because the earth was the centre of the universe and all the planets rotated around it in circles. The Ptolemaic system was based not only on astronomical observations but also on some basic aspects of perception. If the earth was moving would it not be felt as such? However, we are not aware of any such motions and so it was a reasonable inference that the earth-bound observer was stationary and that Ptolemy's eye was at the centre of the universe. Thus, the system was based in part on astronomy and in part on perception.

Ptolemy's harmony of the spheres was disturbed by Nicolaus Copernicus (1473–1543) who proposed that the sun rather than earth was the centre of the universe. That is, although the eye of the observer is located on earth it does not lie at the centre of the world: "Finally we shall place the Sun himself at the centre of the Universe. All this is suggested by the systematic procession of events and

the harmony of the whole Universe, if only we face the facts, as they say, 'with both eyes open'". The eyes of Copernicus were based on earth but the sun is at the centre of this heliocentric system. The new cosmology was based on physics rather than perception. In what follows, we will discuss frames of reference in the context of perception rather than physics. That is, the spatial characteristics of direction, distance and motion will be described with respect to the perceptions of observers themselves rather than to an external origin. The distinction is between subjective (perceived) attributes and objective (physical) measurement. While this distinction is helpful for understanding perception, it is worth noting that 'physical' measurements are themselves the result of applying explicit frames of reference, which are created by human observers and therefore perhaps best described as cognitive rather than objective. While frames of reference are made explicit in the cognitions of physics, they are not so in our perceptions and must be inferred from experimental findings.

The difficulty in separating these two forms of representing space is evident when considering perceived distance. Distinguishing distance as perceived by an observer from that measured by some physical instrument (like a tape measure) is an issue that has taxed writers on perception and for which there is no clear nomenclature. We will return to this issue later in this chapter.

Frames of reference

Our language is replete with terms relating to where things are with respect to one another. For example, is this book on the table or beneath the light or beside the window or above the floor? All these questions refer to relations between objects. In these statements there is an implicit assumption of the greater stability of one of the objects (table, light, window, floor) relative to the book. Some objects may be considered to be more stable than others, but those objects could be changing place, too. If the book is being read on a train then all the objects would be moving in the same direction. The statements above would still apply if the carriage was taken as a frame of reference, but the carriage (and its contents) could be located with respect to a larger frame of reference, like the surface of the earth.

A physicist would describe the locations of objects in this way, and also draw attention to the fact that the earth is moving relative to the sun, which is not, of course, stationary with respect to the centre of the Milky Way and so on. A physicist assumes that these nested frames of reference can be defined in abstract terms; the descriptions do not make any reference to the location of an observer. In our example, the reader of the book can change location with respect both to it and to other frames of reference like the floor. Physicists relate objects in the world to other objects; they are not (as physicists) a part of that world themselves. If we are to discuss perceived location then we must also take the perceiver into account. Thus the situation is somewhat different, and more complex, for perception than for physics. In practice, there is a point at which the frame of reference is sufficiently broad to encompass perception and action. This is typically the surface of the earth, which is perceived to be stationary.

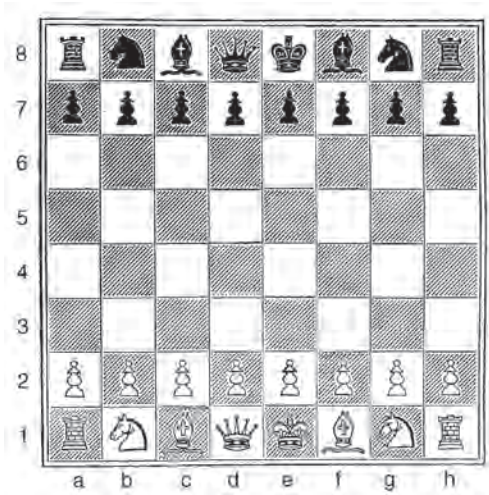


Figure 6.1 The layout of a chessboard with symbols for rows and columns, so that every square can be specified.

As we shall see, under certain conditions this assumption can lead to errors in perception of self location.

The concept of frames of reference is, accordingly, of critical importance in understanding the issues in this and subsequent chapters. By way of analogy, Locke's chessboard could be likened to the retina, and the location of a piece could correspond to a particular point on the retina. If the chessboard (or retina) is adopted as a frame of reference then the location of a piece with respect to it would be called, in our terminology, *retinocentric*. All that is required to specify a particular square is the adoption of a point on the surface as the origin with specified units of separation. On a chessboard (Figure 6.1) the convention is to adopt the black square with the White Queen's Castle as row *a* column *1* (*a1*) and then the rest of the 64 squares can be given specific locations (like *b6*). This is essentially how complex arrays (like pixels on a computer monitor) can be specified: every pixel can be defined with respect to some origin. Usually, the arrays are arranged like a larger version of the chessboard; rather than 64 possible positions there are many thousands each of which can be assigned a horizontal and vertical value, known as a 'local sign'. In the case of the retina it is likely that the origin would correspond to the fovea in the centre rather than an extremity as in the chessboard example.

In a historical sense, the concept of location on the retina has played an important role in theories of space perception. It was called the local sign by the philosopher Rudolf Hermann Lotze (1817–1881) and it was closely linked to the perception of direction. Hering utilised the concept of local signs in his disputes with Helmholtz over stereoscopic vision (see Chapter 3). The stimulation of a point on the retina delivers a retinocentric (or local sign) signal that defines its location with respect to the coordinates of the retinal surface. That is, its value

with respect to an origin in the retina itself. Possible coordinate systems that operate within the retina will be described in the next section.

However, the situation is a little more complex for vision because we have two retinas: it is as though we had two similar chessboards, side-by-side; which one should we use? That is, a single object in space can stimulate points on each eye, so that there are two origins and two sets of coordinate values. Despite this, our perceptual experience is not in duplicate: we see objects singly with two eyes and in a single direction. Moreover, this direction does not correspond to that from one or other eye but from a point between them, often termed the egocentre or cyclopean eye. We therefore refer to this as the *egocentric* frame of reference and it provides a single origin from which the direction of objects can be assigned. It is as if we have derived a line of direction in which objects are seen, but it does not enable us to determine how far along the line they are positioned. A large part of visual processing in the brain is concerned with combining the two retinocentric signals to derive a single egocentric origin relative to which directions can be assigned.

The term 'egocentric' is now in wide use in areas like spatial memory, where objects can be remembered with respect to the observer. However, this usage is rather different from the tighter definition that we apply to perception, as will be made apparent in what follows. The egocentric reference enables us to determine the directions in which objects appear with respect to a point located between the eyes. We are able to judge directions with respect to the egocentre, so that we can point to them. It is like seeing that an object is in front of us but without seeing how far away it is. In order to locate objects in space we need to have some indication of distance as well as direction, and this involves adopting yet another frame of reference, which we have called *geocentric*. It is three-dimensional, and it can be used to guide our actions. Distance cannot be egocentric, because it refers to a measurement in a frame of reference external to the observer. If you are asked to say in which direction the book you are reading lies, then that constitutes an egocentric judgement. If you are asked to say how far away it is from you, then your answer will be geocentric – with respect to a frame of reference provided by the world around you, in which you are yourself located. The fact that you are reporting the distance of the book from you does not make the judgement 'egocentric'. That terminology has often been adopted, but it is clearly inaccurate. There is no generally accepted alternative terminology, but it is important to avoid the confounding of a frame of reference (whether defined physically or represented in perception) with the type of report that is required from an observer. In short, the egocentric reference enables perception of the direction of objects with respect to the observer and a geocentric reference involves the distance objects are away along the line of egocentric direction.

The geocentric reference is not, of course, the final frame of reference that could be adopted, but it is the one that is both relevant and vital for our behaviour and survival. We know that the earth rotates around the sun, but our experience each day is that the sun rises and appears to rotate around the earth. When we need to use larger frames of reference (as in navigation by air and sea) we tend to use

instruments or memory rather than rely on the senses alone because there have been no evolutionary pressures for us to develop detectors for them.

The use of retinocentric, egocentric, and geocentric frames of reference enables us to discuss the perception of location in terms of natural articulations of the body – the stimulation of the retina, the rotation of the eyes and the motion of the head. However, Locke's example of spatial relativity was based on a single chess piece. Suppose that there are several pieces on the chessboard. Each could be specified with respect to the origin (equivalent to retinocentric coordinates), but they could also be described with respect to one another. To state that the King's pawn can take the Black Bishop (or that the Black Bishop can take the King's pawn) does not indicate where the two pieces are on the board, but only their relative locations. Similarly, it is possible to determine retinal locations in this way, so that they relate to the particular pattern of stimulation. Certain motion phenomena can best be understood by introducing such a frame of reference, which we have called *patterncentric*. One potentially useful feature of patterncentric information is that it does not change as the eyes move in the head; the relative position of two stimuli on the retina remains the same even if the retina moves. Under normal circumstances the patterncentric frame of reference will correspond to the geocentric, but there are circumstances under which they are dissociated. We can learn a lot about spatial vision by studying these, for example when two objects are at different distances from the observer and the retina is displaced through self-movement, or with suitable experimental methods, as described below.

The concept of a patterncentric frame of reference can be illustrated by means of an example. In Chapters 2 and 5 we described the motion aftereffect, an illusion in which the observer watches a surface (e.g. a waterfall) moving in one direction for 30 seconds or more (adaptation). If it stops moving, or if the observer looks at another stationary surface (like the rocks surrounding the waterfall), then the observer sees illusory motion in a direction opposite to the previously observed movement. The illusion will be further discussed in Chapter 7, along with other motion phenomena. However, the authors' work on the motion aftereffect has concerned the role of patterncentric and other frames of reference in inducing and observing the illusion, and description of a typical experiment may therefore clarify key concepts.

The motion aftereffect is normally observed under conditions that make it impossible to separate the frames of reference in operation. Thus, an image moves with respect to the retina (retinocentric), but the eyes and/or head may move, yielding a different geocentric motion. There are also visible features other than the moving surface, so motion is also with respect to them (patterncentric). We can also ask whether the frame of reference necessary for inducing the illusion is the same as that required to elicit it subsequently. In the case of a waterfall, it would not be possible to make the rocks move upwards and the water remain static. However, within the confines of a laboratory it is possible to do something similar, if all the visible features are under experimental control. In order to address these issues, we devised a method for studying the phenomenon that allowed the frames of reference to be explicitly defined and controlled.

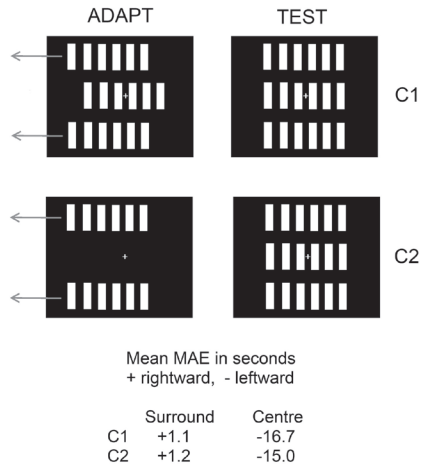


Figure 6.2 Schematic illustration of the conditions (C1 and C2) of the motion aftereffect (MAE) experiment described in the text, with a summary of the main results. The grating bars are shown white for clarity, but were low contrast to ensure that nothing else was visible in the testing room. During the Adapt stage, the upper and lower gratings moved towards the left (indicated by the arrows) and then reappeared at the starting position before another cycle of movement. Observers adapted for ten cycles and then pressed a key whenever they saw a motion aftereffect in the stationary surround gratings or the central grating of the Test display. The results showed clearly that the motion aftereffect was effectively absent from the surround gratings, even though they moved. By contrast, it was strong in the central grating, even though there was no central grating during adaptation in Condition 2.

Observers sat in a completely dark room, in which nothing was visible except the low contrast visual stimulus. The stimulus was typically one or more strips of grey grating bars, as illustrated in Figure 6.2. A grating strip moved from left to right, as a block. It was then blanked and restarted from its original position. It never reached the edge of the display screen, so there was no appearance and disappearance of the grating at its edges. The only patterncentric information was therefore the motion of one strip with respect to another. Since the observers maintained fixation, retinocentric motion was unaffected by eye or head movements. One experiment, whose conditions and results are summarised in Figure 6.2, can serve as an example of the methodology and the type of experimental question that could be addressed. We wanted to know whether the frame of reference for motion during the initial adaptation phase was the same as that needed to elicit the illusion during testing. In Condition 1 (C1 in Figure 6.2) the upper and lower surround gratings moved repeatedly from right to left and the central grating was stationary. In Condition 2 (C2) the central grating was absent. The duration of the motion aftereffect was tested for both conditions with all three gratings present in a stationary aligned position. The results were very clear and are summarised in Figure 6.2. There was almost no motion seen in the upper and lower gratings, which had actually moved, both physically and

with respect to the retina. However, the central grating was reported as having a large motion aftereffect, even though it had been stationary in C1 and absent altogether in C2. This motion aftereffect was in the same direction as the adapting motion of the surround gratings. The surprising conclusion from this and other experiments is that retinocentric motion alone is sufficient for adaptation. There do not need to be other visible visual features to define patterncentric motion. However, this adaptation is only expressed, and therefore able to be seen and measured as a motion aftereffect, if there are visible features both where the adapting movement occurred and in areas not previously stimulated by motion. In summary we concluded that retinocentric motion is sufficient for adaptation but patterncentric stimulation is necessary for expressing the motion aftereffect. This finding has implications for the mechanisms underlying the phenomenon and makes a simple explanation in terms of the adaptation of retinocentric motion detectors untenable. The methodology shows that it is both possible and important to specify and control the operative frames of reference in experiments on motion perception. This conclusion can be generalised to the perception of orientation and direction, where the perceptual frames of reference are similarly critical but very often ignored.

While our perception of space is based upon these frames of reference, our navigation through space might be more complex. For example, if you describe the room you are in while reading this book, then the relative locations of objects that are not visible can be detailed. Moreover, your actions can be planned and executed with respect to these unseen but remembered objects. This is an aspect of memory that is called spatial cognition and similar terms have been adopted to those we use. Unfortunately, they do not have the same meaning, which can be a source of confusion. Theories of spatial cognition draw a distinction between egocentric and allocentric (sometimes called geocentric) representations of space – between aspects that are available to the viewers and those that are independent of them. For example, the location of the text you are reading can be specified perceptually, but the objects behind you cannot. The conceptual confusion is caused by the manner in which egocentric and allocentric frames of reference are defined in spatial cognition. We define egocentric as having an origin in the observer but yielding only directional (angular) values; in order to locate objects distances require to be allocated to directions, yielding a geocentric representation. Yet it is precisely this geocentric representation that is called 'egocentric' in spatial cognition. To add further to the confusion, the term 'allocentric' is often referred to as 'geocentric'. In our terminology, geocentric representation is viewer centred and assigns distances to egocentric directions: perception is geocentric. Cognition may well be allocentric, but it is a concept that we consider beyond the scope of this book. Accordingly, we will retain the more tightly defined frames of reference for perception that we introduced in the first edition of the book – retinocentric, patterncentric, egocentric and geocentric.

In summary, the retinocentric frame of reference relates to location on the retina, the fovea can be considered as the origin from which coordinates are assigned. The patterncentric frame of reference can take any point on the retina

as an (arbitrary) origin and assign the coordinates of another point or points with respect to it. The egocentric reference frame combines signals from the two eyes to define a single origin, which is typically between the eyes; directions are assigned with respect to this egocentric origin. The geocentric frame of reference combines distance with direction so that objects can be located in three-dimensional space; the origin remains within the observer. The dimensions used at the geocentric level are those that apply to three-dimensional space. This enables us to perceive space three-dimensionally and to control our actions in it. Unlike physical descriptions of space, which are viewpoint free, we carry the origin of our geocentric representations around with us.

The discussions of location, motion, and recognition in this and the following chapters will all be couched in terms of these four frames of reference as we believe that they provide a logical basis for understanding vision. Frames of reference are being incorporated explicitly in many theories of perception, but neither the terminology nor the concepts themselves are equivalent. This applies particularly to theories of spatial cognition and to object recognition.

Coordinate systems

In order to use the information available within a frame of reference we require a defined origin and also units which determine the separation of points from the origin. That is, a coordinate system is required. The one with which we are perhaps most familiar is the Cartesian system for drawing graphs; it uses X and Y axes, which are usually represented as horizontal and vertical, respectively. Such a graph is two-dimensional, one for breadth (X) and the other for height (Y). If a third dimension is added at right angles to the other two then depth can be specified along the Z axis. The description made above in terms of a chessboard can be amplified with respect to points drawn on a two-dimensional graph. Figure 6.3a shows a graph with the X and Y axes marked. Any point within the graph can be specified; here the point is at values X = 2 and Y = 2, and it is marked by a circle. The location and intersection of the coordinate axes define a frame of

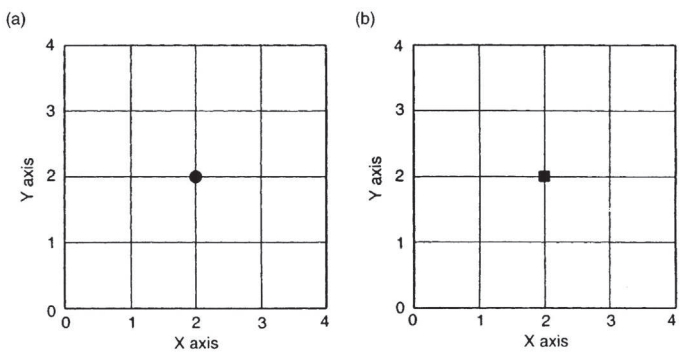


Figure 6.3 Two graphs with circular and square symbols at the same X and Y coordinate values (2, 2).

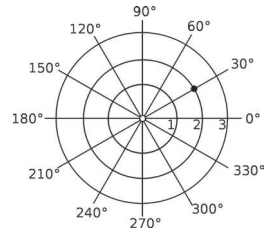


Figure 6.4 Polar coordinates in which the origin is at the centre and any point can be specified by the distance from the centre (r) and the angle in an anticlockwise direction from the horizontal (θ). The black dot has coordinate values of (2, 30 deg).

reference relative to which the point is located. Within the frame of reference defined by this origin and these axes, nothing can be said about any point outside them: only values within these coordinates can be give with respect to this origin. Figure 6.3b represents a similar graph with the point (marked by a square) at the same coordinate values (X = 2 and Y = 2) with respect to the origin. If we take the axes as the only frame of reference in each case then the points have the same locations relative to their origins because they have the same coordinates on axes of the same scale.

An alternative coordinate system is based on angles and directions with respect to a fixed origin. An example of such polar coordinates is shown in Figure 6.4. The origin (the white dot in the centre) is specified and a horizontal, rightward direction (a vector) from it is assigned a value of 0 deg, and units are applied to this extension. The location of the point shown is then defined by the distance along the projecting line (*r*) and the angular deviation from the horizontal (θ). Angular values are given as rotations in an anticlockwise direction. Thus, the black dot in Figure 6.4 would have coordinate values of (2, 30 deg) and, as with Cartesian coordinates, any point within the range of the graph can be given a unique pair of coordinates (*r*, θ). However, for convenience of description we will use the Cartesian system in what follows, although the same arguments can be made using either system.

One virtue of the polar coordinate system is that it has a superficial similarity to the retina, as the origin could be assigned to the fovea. However, as was noted in Chapter 4, the retina is not uniformly covered with receptors and the types of receptors themselves differ. Cones are found exclusively in the fovea and their density declines rapidly with retinal eccentricity so that rods predominate and are found exclusively in the periphery. Thus, graphical coordinates retain a uniformity of their units which does not apply to the eye, and this point should be kept in mind.

Another aspect of polar coordinates that maps on to our discussion of frames of reference concerns the distinction between direction and distance. Suppose you are looking at an observer from above, so that the origin is the egocentre and the straight ahead is at 0 deg. The direction of the object symbolised by the black dot is at 30 deg. That is, the object lies on the line at 30 deg to the straight

ahead, but there is no indication of how far along the line it is. In order to locate the object its distance as well as its direction needs to be specified. The direction in which the object lies is egocentric so that only the angular deviation from the reference (straight ahead) is available. By specifying the distance as well as the direction the frame of reference is geocentric. This would allow the observer to behave appropriately with respect to the object, for example by reaching for it and grasping it.

Hierarchies of coordinate systems

In the context of Cartesian coordinates, the intersection of the X and Y axes defines a point (the origin) relative to which all other points are plotted. It is only when a further frame of reference is adopted that a measure of the enclosed one can be taken. For instance, the graphs represented in Figure 6.3 are positioned differently on the printed page: if the page is taken as a frame of reference then the locations of the two graphs on it can be measured, and differences between them can be described. Therefore, taking the page as the frame of reference the circle and square symbols have different locations, even though their coordinate values (2, 2) are the same with respect to their origins. The page itself can be in different positions, but this can only be measured if a larger frame of reference still is adopted, like the table, the ceiling, or the surface of the earth. In such a descriptive system no information about locations with respect to one frame of reference is lost by moving up to another one; rather information that was not available is introduced by adopting a superordinate frame of reference. Again, this is like considering the chessboard, the cabin, and the land as frames of reference in Locke's example.

Statements about the location of objects are restricted to the highest frame of reference in operation. For example, at the level of the graph in Figure 6.3a, the coordinates of the circular symbol can be specified, but no statements can be made about the location of the axes, because they define the frame of reference relative to which all other points are specified. It is only when a higher frame of reference is adopted, like the page, that the location of the axes can be described, and the location of the page can only be specified if we consider a frame of reference relative to which it can be located, like the surface of the earth.

The graphical example is conceptually similar to what is happening in spatial perception, because the perceiver has many articulating parts: the eye can move in the orbit, one eye can move with respect to the other, the head can move on the shoulders and the body on the ground. The retina in one eye is somewhat like the drawn X and Y axes on the graph paper; different points on the retinal surface can be stimulated, but they can only be measured with respect to the retina itself. The second sheet of graph paper is like the other eye. The information about locations on each eye could be derived by adopting a similar frame of reference for both, like the coordinate axes of the two graphs. In the example that we have drawn in Figure 6.3, the coordinates on each graph are (2, 2) and (2, 2), and the units of length are the same in both. If the axes of the two graphs were superimposed on

one another, then the points would be in correspondence. The page can be likened to the head; whenever the page moves the two graphs move with it. Similarly, whenever the head moves the two eyes are displaced in space. Therefore, changes in retinal stimulation can be produced by movements of the eye with a fixed head, by movements of the head without eye movements, or by a combination of the two. The changes in position of the page (or head) occur with respect to a larger frame of reference, namely the surface of the earth. For all practical purposes, the surface of the earth will be taken as the final frame of reference and will accordingly be defined as stationary.

Visual orientation

Locke's chessboard example implicitly involved another aspect of spatial perception that is so pervasive that it is often overlooked: orientation. Gravity was operating in all the conditions Locke mentioned and it provides a frame of reference relative to which we and objects are aligned. Unless we are in free fall, gravity will be a force operating on us all the time and sensory systems have evolved to detect its direction with precision. The vestibular system is located in the inner ear and consists of two components which respond to different accelerations – linear and angular. The otolith organs respond to linear accelerations, of which gravity is the most constant; the semicircular canals respond to angular accelerations, like those involved in turning the head. Thus, the issues raised about frames of reference, and coordinate systems generally, can be illustrated by examples from visual orientation. Moreover, these considerations make evident the extent to which vision is integrated with other sensory systems that signal vital spatial information.

The perception of object orientation is an essential requirement for appropriate behaviour in three-dimensional space. Whenever we grasp an object we need to determine its global orientation with respect to the surface supporting it, and those of it parts. Object orientation is initially registered via systems (the eyes and the head) which do not remain in a constant alignment with respect to gravity, but we are still able to grasp objects appropriately despite large variations in the location and orientation of the head. Thus, the representation of orientation in vision is likely to involve information in addition to that provided by light and it can be examined with respect to the four frames of reference mentioned above: retinocentric, patterncentric, egocentric and geocentric.

The great strides made in uncovering the neurophysiological processes involved in the early registration of orientation were described in Chapter 5: single cells in the visual cortex of higher mammals are selectively tuned for orientation, and other stimulus dimensions, like motion and colour, are processed with respect to orientation. It is not, however, straightforward to allocate even the simplest aspects of visual orientation to the operation of physiological orientation detectors, because the light detectors are retinocentric whereas visual orientation is geocentric. That is, the orientation detectors are coded with respect to retinal coordinates, while visual orientation is generally in terms of the gravitational

reference. There requires to be some way in which orientation, coded initially in retinocentric coordinates, can be assigned to gravitational coordinates; this could be achieved by processing with respect to intervening patterncentric and egocentric frames of reference. It is assumed that all perception occurs at the geocentric level, but a given phenomenon might reflect the operation of lower levels if the signals from that level are not modified at subsequent levels. Here visual orientation will be examined initially in the context of tasks devoid of any visual reference cues (like an isolated light line in darkness) before introducing the additional influence of patterncentric processes.

Visual orientation tasks often examine only a very circumscribed set of orientations – typically vertical and/or horizontal – which are defined with respect to an external reference: the direction of gravity. Many experiments have shown that the vertical and horizontal orientations can be perceived with great accuracy and precision even in the absence of any visual reference: a line of light in an otherwise dark room can be adjusted by an upright observer to within 1 deg of the gravitational vertical. Such precision might be taken to reflect the operation of cortical orientation detectors alone, but this cannot readily be decided from experiments involving upright observers: only when retinal and gravitational orientations are dissociated can their respective contributions be determined.

Frames of reference for visual orientation

Orientation is a relative term, and so it is necessary to specify the reference relative to which it is registered. In physics the reference is gravity, which can readily be determined by suspending a plumb bob and noting the direction in which it hangs. In perception there are several nested frames of reference. The initial one is retinocentric: that is, orientation is coded in terms of retinal coordinates. At the retinocentric level, orientation can only be signalled relative to the retinal coordinates stimulated. Thus, there would be no way of distinguishing between the inclination of a line and the inclination of the eye: the same retinal coordinates would be stimulated by a line in the gravitational vertical for an upright eye and by a horizontal line for a horizontal eye. The representation of retinocentric orientation is not confined to the retina; it can be retained at the level of the cortex if orientation is still specified in terms of retinal coordinates. Indeed, the evidence described in Chapter 5 indicates that the majority of cortical orientation detectors are retinocentric – they would respond similarly to a line at a given retinal orientation no matter what the position of the eye.

The retinocentric signals from each eye are combined to define an egocentric frame of reference. At this level there is a single signal for orientation with respect to an origin at the egocentre (or cyclopean eye) with the reference axis as the head median plane. This is an important stage of integration because orientation is assigned a single value that is tied to the orientation of the head. At the egocentric level, orientation can only be signalled relative to the head coordinates; thus, there would be no way of distinguishing between the inclination of an object and the inclination of the head. The final stage incorporates signals from postural

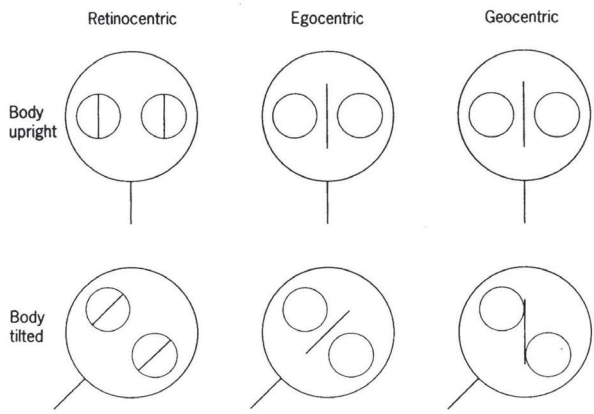


Figure 6.5. Schematic representation of the eyes (small circles), head (large circle) and body (line) with respect to retinocentric, egocentric and geocentric frames of reference for an upright and a tilted head and body.

systems (like the otolith organs signalling the direction of gravity) which allocate orientation to a geocentric frame of reference; that is, with the egocentre as the origin and gravity as the reference axis. The retinocentric, egocentric and geocentric frames of reference are shown schematically in Figure 6.5, both for an upright and an inclined observer; in each case the head and body are aligned so that body tilt is the same as head tilt. It is clear that egocentric and geocentric orientations cannot be distinguished when the body is upright. However, when the body is tilted the difference is plain to see. You can readily demonstrate the geocentric nature of visual orientation by looking straight ahead and tilting the head sideways: does the world appear tilted or upright? When the head is upright, vertical lines in the world will stimulate retinal coordinates that are vertical but this does not apply when the head it tilted.

Experiments with observers viewing a line of light in an otherwise dark room support these outcomes. However, we are seldom in such constrained environments; we judge visual orientation with respect to visually rich contexts. Moreover, many of the objects in our visual environment provide clues to the direction of gravity as they have been constructed within the constraints of gravity. The rooms through which we move are likely to be rectangular with vertical walls and horizontal floors (although these rules are not always followed in psychological laboratories!). Accordingly, in addition to the three frames of reference shown in Figure 6.5 another, patterncentric reference is introduced when a variety of orientations is present simultaneously. It is not easy to separate the patterncentric cues to vertical in a natural environment, but it could easily be done in the laboratory. For example, if the majority of contours are in the same orientation (and inclined with respect to the retinal meridian) then they can establish a patterncentric reference relative to which other orientations can be assigned. The patterncentric frame of reference would not then correspond to that

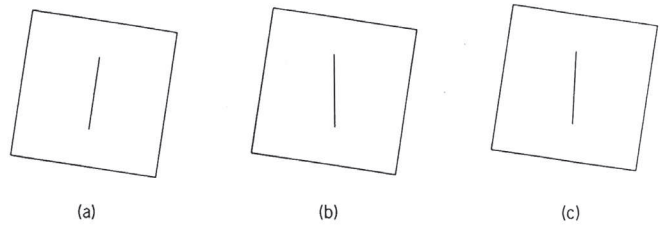


Figure 6.6 Schematic representation of a tilted room enclosing a line, with the latter aligned with a) the patterncentric and b) the geocentric frames of reference, and c) a compromise between the two.

defined retinocentrically, and some resolution of this discrepancy is required. This is displayed schematically in Figure 6.6 in the context of a tilted square (which might be thought of as a room). The principal axes of the square would define a patterncentric frame of reference. If this was operating in isolation it would determine the visual orientation of other contours enclosed within it, so that a line would appear vertical when parallel to the sides of the tilted room (Figure 6.6a). If it had no effect then the line would look vertical when in alignment with gravity (Figure 6.6b). In practice, a compromise is reached between the patterncentric and geocentric frames of reference, as indicated in Figure 6.6c.

Retinocentric orientation

There must be some retinal local signs or mapping coordinates to determine retinocentric orientation. This could be in terms of polar coordinates with the fovea as origin and the retinal meridian as a cardinal axis (see Figure 6.4). The retinal vertical would correspond to the line passing through the origin along the 90/270 deg axis. However, the eye can rotate around the axis perpendicular to the page and through the origin (around the Z axis). These are called torsional eye movements and are relevant to specifying retinocentric orientation. The retinal meridian and gravitational vertical correspond when the eye is in the primary position; this is when the head is upright with the eyes parallel and directed to the horizon. Torsional eye position is affected by stimulation of the otolith organs, which are the principal source of information for allocating orientation with respect to gravity both when upright and during lateral tilt.

One method of determining whether visual orientation corresponds to retinocentric orientation in an upright observer is to measure the former when the eye is inclined. This occurs when the gaze is directed obliquely; the eye undergoes torsion so that the retinal meridian is no longer aligned with gravity. Under these circumstances the visual vertical is also inclined with respect to gravity, but not by as much as the retinal meridian. When the eyes of an upright observer are in the primary position they are normally aligned closely to the direction of gravity. This situation is markedly disrupted following vestibular surgery when both eyes assume many degrees of torsion. Measures of visual and eye orientation

following vestibular surgery point to a close association between visual and retinal orientations, but not their equation. In the absence of strong evidence to the contrary, it seems reasonable to accept the assumption that muscular signals for eye orientation do not influence visual orientation.

Egocentric orientation

The combination of the two retinocentric signals determines the egocentric frame of reference. That is, an origin is taken at the egocentre and orientation is registered with respect to that. Egocentric visual orientation refers to judgements made of the axis of the head. This is called the head median plane, and can be thought of as the plane passing through the egocentre and bisecting the head. If the head and body are aligned then the head median plane corresponds to that of the whole body. When the head and body are upright the direction of gravity corresponds to the head median plane but they can be dissociated by tilting the head or whole body (as shown schematically in Figure 6.5). Under these conditions visual vertical judgements remain aligned approximately to the direction of gravity whereas those for the head axis do not. However, when the head is tilted laterally with respect to gravity the eyes do not remain aligned with the head: they undergo rotation in the opposite direction to tilt (called countertorsion or counterrolling) that is about 10 per cent of the degree of tilt. For example, if the head was tilted 30 deg from the vertical the eyes would undergo countertorsion of about 3 deg, so the eyes would be inclined around 27 deg from vertical. Some experiments have involved inclination of the whole body whereas other have tilted the head alone relative to an upright body. Judgements of egocentric orientation during body tilt show slight variations from the body median plane, deviating more than body tilt for small tilts and less for larger tilts. The situation for head tilt alone seems rather clearer: the orientation of the visual head axis is inclined in the opposite direction to that of head tilt by an amount that increases with head tilt and that approximates the degree of countertorsion operating. Unlike measures of the visual vertical, the action of vestibular input can be excluded by examining judgements of the visual head median plane when the head is inclined laterally with respect to a supine body. That is, observers lie on their backs and view a line of light directly above them; their task is to align the line with the head median plane when the head is aligned with the body or tilted laterally to it. Under these conditions the visual head median plane displays a different function to that with an upright body: judgements are approximately aligned with the head axis when supine but are inclined in the opposite direction to head tilt when upright.

Geocentric orientation

The visual vertical is located relatively close to the gravitational vertical during lateral head or body tilt from a normally upright position. There are, however, systematic deviations depending upon the degree of tilt. In 1861, Hermann Aubert (1826–1892) described how a gravitationally vertical line appeared to be tilted in

the opposite direction to the body for large tilts, and later, in 1916, Georg Elias Müller (1850–1934) reported the opposite for small tilts (less than 60 deg). That is, the visual vertical is displaced in the same direction of the gravitational vertical for body tilts above about 60 deg, and in the opposite direction for smaller tilts: these are referred to as the A- and E-effects, respectively. More attention has been focused on these deviations from the gravitational vertical than on the much greater inclinations with respect to the retinocentric and egocentric reference axes. The principal source for transforming the egocentric to geocentric orientation is undoubtedly the otolith organs of the vestibular system: individuals lacking vestibular function locate the visual vertical much closer to the egocentric axis during tilt and centrifugation than do normals, as do patients following surgical section of the vestibular nerve.

Thus, under normal conditions our perception of orientation is geocentric. The direction of gravity also corresponds to the main lines visible in our environment and so patterncentric orientations are aligned with gravity. Nonetheless, experiments with isolated light lines in otherwise dark spaces indicate that we continue to perceive orientation geocentrically without the aid of patterncentric cues.

Patterncentric orientation

Gravitational and patterncentric orientations can occasionally be dissociated by natural disasters, like earthquakes. Following an earthquake some years ago, many of the buildings were inclined by several degrees. One consequence of this mismatch between the predominant visual orientation and gravity was an increase in referrals to doctors of problems associated with balance and posture! As will be described below, a less intrusive way of separating the patterncentric frame from the direction of gravity is to tilt a room with respect to an observer.

Patterncentric orientation has received most attention in the context of orientation illusions, which can be taken as instances of patterncentric interactions. That is, when the presence of one orientation influences the perception of another for an upright observer. Here, we will consider those instances in which the pattern elements do not intersect, although many orientation illusions do involve such figures. When a tilted annular grating surrounds a vertical grating, the latter appears tilted in the opposite direction (see Figure 5.21); this tilt illusion is about 2 deg when the surround is inclined by 10–15 deg, and it is dependent upon the spatial frequencies of the two gratings. Similar shifts in visual orientation can be produced by enclosing a single line within an inclined outline rectangle; this is called the rod-and-frame illusion (shown in Figure 6.6), and its maximum magnitude is about 2 deg for surround orientations of 15 deg. The patterncentric frame of reference reflects relational orientations within a display and it is expressed in relational rather than retinocentric terms. The two signals are extracted in parallel and they are combined at the egocentric level.

The rod-and-frame illusion is a simplification of a more elaborate arrangement in which a well-illuminated room is inclined laterally with respect to an upright

Figure 6.7 (a) A photograph of a tilted room in which an upright person is standing; the photograph has been orientated so that the main lines of the room correspond to those of the page. (b) The upright person in the tilted room, with the photograph orientated so that the person is aligned with the sides of the page.

observer. Under these conditions the visual vertical is displaced in the direction of the tilted room. These orientation shifts are more compelling than the tilt illusions above, and also influence the perception of body position: the upright body feels inclined with respect to gravity, and apparently upright when inclined in the direction of room tilt. Fairground illusionists appreciated this long ago, and tilted rooms were constructed to create disorientation and amusement in their customers. Figure 6.7a shows one of the authors standing upright in a tilted room. Thus, there is visual-vestibular conflict in a manner not found with tilt illusions. That is, the retinocentric and patterncentric signals combine as for the rod-and-frame illusion, with the patterncentric input also modifying the value of perceived body position.

Presenting a picture of this situation, as in Figure 6.7a, is itself confusing because the sides of the book are also providing a patterncentric frame of reference so that the sides of the pictured room are parallel with those of the page. A patterncentric conflict is induced if the picture is rotated so that the person is aligned with the page but the pictured room is not (Figure 6.7b). In this context, it is worthwhile thinking what is meant by the term ‘upside down’ when you are looking at a picture in a book or on a screen.

Visual direction

Visual direction can be considered as a line extending from a point in the perceiver. As was the case for visual orientation, we will start our treatment of visual direction by considering simplified situations, like a stationary observer. This has the advantage of building up a scheme of spatial representation in a systematic

way. The information available at early stages (like the retina) is generally inadequate for guiding behaviour, and it needs to be transformed successively by other sources of information (like eye and body positions) in order to achieve a representation of three-dimensional space.

Retinocentric direction

Suppose a point of light in an otherwise dark room is observed by a single, stationary eye. How can observers determine its direction? One of the sources of information that must be used is the location on the retina that is stimulated. This is referred to as a local sign. Essentially, a local sign can be considered as a two-dimensional coordinate reference on the retina itself. It is likely that the fovea acts as the origin for the retinal frame of reference, and positions with respect to it can be discussed in terms of the familiar Cartesian coordinate system (see Figure 6.8), although other coordinate systems (like the polar plots of Figure 6.4) could equally well apply. Within this retinocentric frame of reference there is no information about the position of the retina itself, only about points on it. If one eye moved slightly, as is illustrated schematically in Figure 6.9, then different retinal coordinates (or local signs) would be stimulated. At the retinocentric level there would be no means of determining whether this was a consequence of motion of the point or displacement of the eye.

Now consider a situation in which two points of light are present in an otherwise dark room. If the directions of each point are determined independently then their directions with respect to one another could also be computed. This is essentially the comparison of the two retinocentric values (see Figure 6.10). Thus, the coordinates of one point (x1, y1) could be subtracted from those of the second point (x2, y2) to determine their separation. It would also be possible to derive the relational directions of the two points without determining their retinocentric values: one could be so many retinal units away from the other, so that their separation alone was coded. For example, in Figure 6.10 one point (x1, y1) could be taken as the origin of a coordinate system and the separation from (x2, y2) would be given by its coordinates; alternatively, the second point could act as the origin. Such a relational system can take either point as the frame of reference for

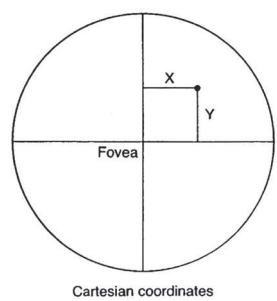


Figure 6.8 Circular schematic of a retina representing a point at a given location in terms of Cartesian coordinates. The fovea is taken as the origin for the coordinate system.

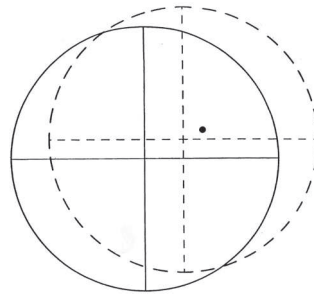


Figure 6.9 A point (the black dot) located at a constant position in space would stimulate different local signs (retinal coordinate values) if the eye moved, as indicated by its position relative to the solid and dashed circles.

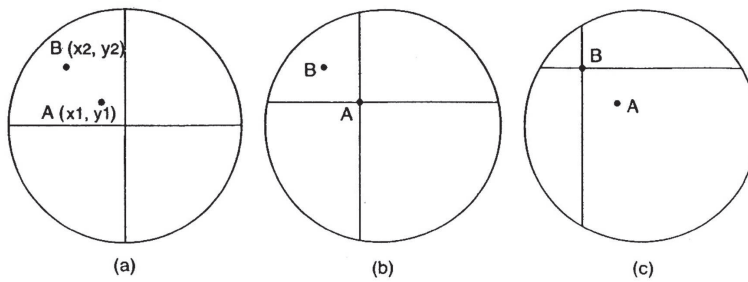


Figure 6.10 (a) Two points (A and B) could be described by their Cartesian coordinates (x1, y1) and (x2, y2), and their locations relative to one another could be determined from the difference between the coordinates. Alternatively, (b) point A could act as an origin and define a patterncentric frame of reference, (c) as could point B.

the other. So, for example, point A could be the origin and point B is separated from it, or vice versa. With just two points there are no advantages in taking A or B as the relational frame of reference. It is difficult to present an illustration of such a patterncentric scheme because the page on which it would be printed provides an implicit frame of reference! That is, the directions of one light source with respect to the other could be described independently of their local sign values (their locations on the retina) – they could be anywhere as long as both were imaged in the eye. The relational values would code separation independently of direction or orientation.

Such a relational system operating on its own could not account for our ability to determine the directions of isolated lights, but we will see later that the joint operation of retinocentric and patterncentric values assists in the interpretation of several phenomena of perceived movement.

When an isolated target is presented to the peripheral part of the retina it initiates a fixation reflex – the eye moves automatically to locate the target on the fovea. Indeed, the adequate perception of the spatial detail is dependent upon light from the object falling on the central region of the eye, because acuity is best in

and around the fovea. Therefore, we need to be able to direct our eyes to objects that require more detailed spatial resolution in order to determine whether they are of interest to us. For this to occur the retinocentric coordinates must be involved in the control of eye movements. For example, the retinocentric coordinates corresponding to an isolated point of light could be (x2, y2), and these values are used to bring the point onto retinocentric coordinates of (0, 0).

We can determine the visual direction of a point falling on the fovea (with retinocentric values of (0, 0)). A different visual direction is perceived if the eyes have to rotate to fixate the point on the fovea, and the retinocentric signals remain (0, 0). Thus, different visual directions can arise from stimulation of the same retinocentric coordinates. Clearly, the difference in visual direction could not be registered at the retinocentric level. However, there are signals for the position of the eye in the orbit, and if these eye position signals are combined with the retinocentric signals then differences in the direction of points could be discriminated. Therefore, visual direction is dependent upon both retinocentric and eye movement information.

Egocentric direction

So far, we have considered situations in which a target is observed with one eye, whereas normal vision is binocular. When we rotate an open eye the closed one rotates, too. This can easily be demonstrated. Close one eye and gently place the index finger of one hand on the upper eyelid. Look at the other outstretched index finger with the open eye, and follow its movements. You will feel the closed eye rotating in its orbit, yoked to the movements of the open eye: the two eyes work as one. This principle is known as Hering's law of equal innervation: "The two eyes are so related to one another that one cannot be moved independently of the other; rather, the musculature of both eyes reacts simultaneously to one and the same impulse of will".

The integration of eye movements between the two eyes affects visual direction, too, because the direction in which a stationary monocular target is seen depends on the position of the occluded eye. When one eye is occluded, it tends to rotate inwards, towards the nose, and this rotation of the closed eye results in changes in the visual direction of the target seen with the open eye, as is illustrated in Figure 6.11. Moreover, when both eyes fixate on a target its visual direction does not correspond to the retinocentric direction of either eye alone, but it assumes an intermediate direction, midway between the eyes. Not only do the two eyes work as one, but the visual direction of a target appears aligned with a single central eye. This position between the eyes is often referred to as the cyclopean eye, as well as the egocentre. This can easily be demonstrated with the aid of a sheet of paper and two coloured pens. Make two marks at one end of the sheet separated by about 60 mm (somewhat less than the distance between the eyes) and join one to a central fixation point at the opposite end using a red pen and the other using a green pen (see Figure 6.12a). Now, place the sheet slightly below eye level and fixate the point at which the lines intersect. Although each line is pointing

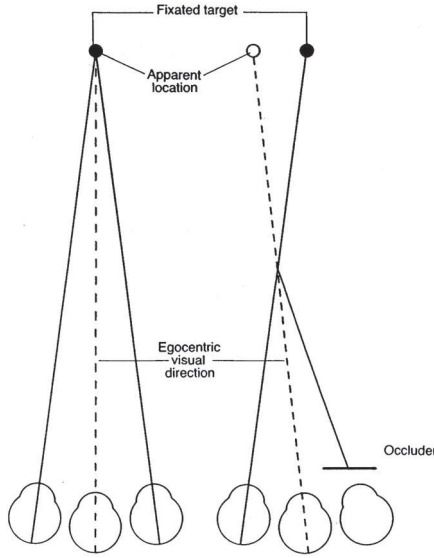


Figure 6.11 The change in visual direction when one eye is occluded. When one eye is occluded it rotates inwards (esophoria), thereby modifying the egocentric visual direction of the target.

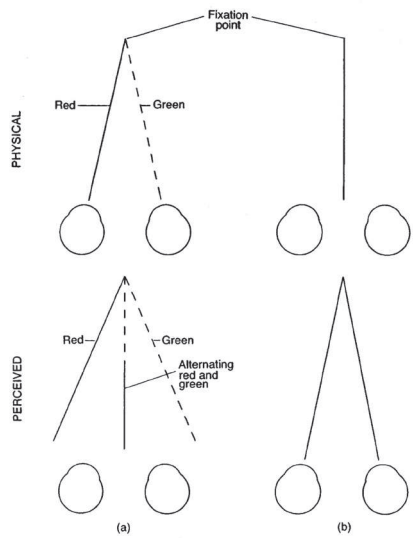


Figure 6.12 (a) The visual direction of red and green lines directed along the optic axes of each eye. (b) The visual direction of a single line in the median plane.

directly at one eye, the impression is of a single line pointing between the eyes (to the cyclopean eye), with two flanking lines pointing wide of the left and right eyes. The central line might also appear to fluctuate between red and green, as a consequence of binocular rivalry.

An even simpler demonstration of the same general point can be made with a single line drawn on a sheet of paper (see Figure 6.12b). If the line is aligned with a point midway between the eyes and the far point is fixated with both eyes then two lines will be visible, one directed at each eye. In this case, the central line appears to be directed to the opposite eye, as can be ascertained by closing the eyes in turn. The initial registration of direction is dependent upon a retinocentric frame of reference, but any judgements made by an observer are dependent on the integrated activity of both eyes. This egocentric frame of reference yields a single visual direction with respect to the egocentre or cyclopean eye. One of the most fundamental aspects of vision is the integration of information from the two eyes to provide a single percept of the world. We have described above how the two sets of eye muscles operate like a single unit; does the same apply to the two sets of retinocentric signals? The evidence suggests that the signals from both retinas are combined independently of those for eye movements. The egocentric frame of reference is based upon information derived from the combination of a binocular retinocentric signal and a binocular eye movement signal.

If the information from each eye is equally weighted, the egocentre is located midway between the two eyes. This is not always the case, because one eye can have a greater influence than the other. For example, if the influence of the right eye is stronger than that of the left then the egocentre would be shifted towards the right eye. It was noted in Chapter 5 that the neural signals from each eye go to adjacent eye preference slabs in the visual cortex, and so the possibility of the eyes contributing differentially to egocentric direction is present in the processing of information in the visual system. In fact one of the oldest tests of eye dominance is based upon assessing visual direction. It was described by Porta in 1593. If you look with both eyes at an object a few metres away and then point at it with your index finger, the finger will appear somewhat blurred but pointing at the target. While keeping the head and finger still, alternately close one eye and then the other. The finger will remain aligned when one eye is open, but not when the other is used. This is a test of sighting eye dominance. It can be refined somewhat by carrying out the task not by pointing a finger but by locating the target in a hole in the centre of a card held in both hands; this avoids problems associated with changing fixation from the far object to the nearer finger, as well as a possible bias from using the preferred hand. About 65 per cent of the population are right eye sighting dominant, 30 per cent are left eye dominant, and the remainder do not have a preferred sighting eye. Porta believed that right-handed people were also right-eyed, but there is no simple link between hand and eye preference.

Binocular single vision

Porta also proposed an intriguingly simple theory to account for binocular single vision. He suggested that we only use one eye at a time! We know that such permanent suppression of signals from one eye does occur in about 4 per cent of the population, usually arising as a consequence of an uncorrected squint in childhood. These individuals can see with each eye alone, although the acuity is much poorer in the deviating eye. When both eyes are open such individuals are functionally monocular, in that neural signals from the normal eye suppress those from the deviating one, as though they only had one eye open. If you happen to be such an individual, it is unlikely to have much effect on your everyday perceptual abilities. A similar form of suppression can be demonstrated in people with normal binocular vision. The vertical and horizontal gratings shown in Figure 6.13 have different visual directions upon initial inspection. However, if you look at the tip of a finger held between the patterns and your eyes its position can be adjusted so that it will appear to point towards the vertical grating when the right eye is open and to the horizontal grating with the left eye. Move the finger until you find a position that is aligned with the base of the centre of the vertical grating with the right eye and the horizontal grating with the left, then keeping your finger as steady as possible fixate its tip with both eyes. Initially the gratings will appear a little blurred, because the state of accommodation is closely linked to the degree of convergence between the eyes. If you persevere for a minute or so, the vertical and horizontal gratings will have the same visual directions but they will not both be visible at the same time: occasionally only one grating will be seen, but more often a complex mosaic made up of local vertical and horizontal regions will be visible. The mosaic is itself dynamic, undergoing constant changes in its composition. This phenomenon, which is called binocular rivalry, is of particular interest because we do not see a stable grid, but rather

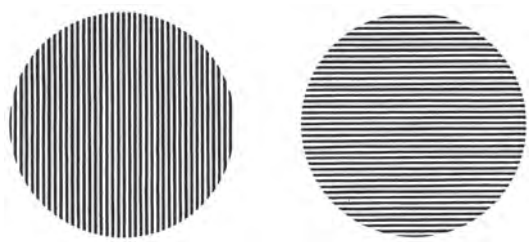


Figure 6.13 Rivalry patterns. Binocular rivalry can be seen with suitable observation of these gratings. If your eyes converge on a point between you and the gratings then the latter can both appear to have the same visual direction. Some people can achieve this without any assistance, for others it is necessary to align, say, a finger-tip with the vertical grating when using the right eye and the horizontal grating when using the left eye. When this is achieved three gratings will be visible – vertical on the left (as seen by the left eye alone), horizontal on the right (as seen by the right eye alone) and the central area where they overlap. Binocular rivalry occurs in this central region, resulting in the fluctuating appearance of the gratings.

an alternation between parts of the two differently oriented gratings. Binocular rivalry also occurs when different colours are presented to each eye – rather than seeing an additive combination of the colours (as would be produced by projecting them independently onto a screen) we see an alternation between them.

The mosaic patterns that we see are made up from local regions in one eye that are suppressing the corresponding regions in the other eye. The fact that such percepts occur at all contradicts Porta's theory, but it does indicate that binocular vision involves aspects of competition as well as cooperation. Competition, or rivalry, takes place when corresponding regions of each eye are presented with different patterns. This provides yet another means for determining eye dominance. If rivalling patterns are viewed for some fixed duration, say 60 seconds, and the duration for which each pattern is completely visible is measured, then the two durations could be compared. The eye receiving the pattern that was seen longer is called rivalry dominant and, surprisingly, it is not necessarily the same as the sighting dominant eye. That is, sighting and rivalry measures reflect different aspects of eye dominance.

The alternative theoretical view concerning binocular single vision to Porta's was proposed at about the same time – at the beginning of the seventeenth century. It was that objects are seen as single when they have the same visual direction, and that visual direction results from stimulation of corresponding points on the two retinas. Corresponding points have the same retinocentric coordinate values with respect to the foveas: if the two retinas were exactly superimposed on top of one another then contacting points would be corresponding. Returning to the example of the graphs mentioned earlier (Figure 6.3), the origins of the two graphs are like the foveas in the two retinas, and the two points on the graphs would be corresponding because they have the same coordinates (2, 2) with respect to the origin. Corresponding points can also be described geometrically: with fixation on a given point, all other points lying on a circle passing through the fixation point and the optical centres of each eye fall on corresponding points. This definition of corresponding points was formulated almost two centuries ago by two German physiologists, Vieth and Müller, and it is now referred to as the Vieth–Müller circle. The geometrical values of the projections to each eye also apply to the egocentre (see Figure 6.14). The geometrical definition is too restrictive in practice; it would predict that only those objects falling on the Vieth–Müller circle would be seen as single and having the same visual directions for each eye.

As the distance of the fixated object from the eyes increases, the Vieth–Müller circle becomes large, and for all intents and purposes it can be thought of as a flat plane at right angles to the straight ahead, rather than a circle. Corresponding points therefore lie on this fronto-parallel plane and the conditions that have been mentioned above approximate to this case. When we fixate on an object in the frontal plane we can still see objects as single when they are slightly in front of or behind it, even though their images do not fall on corresponding retinal locations. This perceived singleness is associated with the perceptual experience of seeing the objects in depth, in front of or behind the point of fixation; that is, the perception corresponds qualitatively with the physical stimulation. However,

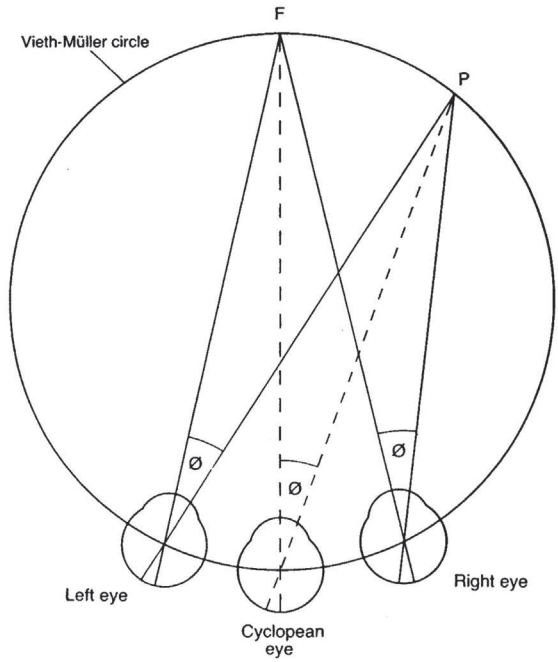


Figure 6.14 Vieth–Müller circle. With binocular fixation on a point (F) a circle can be described that passes through the point and the optical centres of each eye. Any other point (like P) on the circumference of the circle will stimulate corresponding points in each eye. The angle subtended by the two points will be the same at each eye, and also at the cyclopean eye.

the objects seen in depth relative to the point of fixation do not then share the same visual directions for each eye, and neither of these corresponds to the egocentric visual direction. The difference in visual direction from each eye is referred to as retinal disparity. Thus, we see the objects separated in depth with respect to the fixated object, and in directions defined with respect to the egocentre.

There is a limit to the disparity that can be processed to yield depth perception. Only stimuli quite close to the Vieth–Müller circle, in a region referred to as Panum's fusional area, can be seen in depth on the basis of disparity alone. Clearly, much of the field of view will consist of instances where the images of an object fall on non-corresponding areas in each eye, too different to be fused and yield depth. It might be expected that in these cases, objects would be seen as having two visual directions, i.e. seen double. However, we are rarely aware of seeing double images, even though the conditions for their visibility occur constantly. Whenever you fixate on a near object, more distant ones can have radically different alignments for each eye, but they are not often seen double. This can easily be demonstrated by fixating a finger, held about 20 cm from the eyes, and then closing each eye alternately. You will probably be surprised at the large differences in alignment that are evident, and yet they pose few problems for

normal perception. One of the reasons for this is likely to be that the non-fixated objects are slightly out of focus because of the state of accommodation required for viewing the near finger. Another factor is the operation of binocular rivalry: the corresponding regions of each eye will be stimulated by quite different patterns, and so singleness can be served by suppressing one or the other in local regions.

The outcome of the computations based upon patterncentric, retinocentric and eye movement information from both eyes provide a representation of egocentric visual direction. This is a representation of the line along which a single target can be directed or an angle between two targets. The lines or angles do not locate the targets unambiguously because their distances have not been determined. Referring back to Figure 6.4, egocentric direction corresponds to a line at a given angle (here 30 deg) with respect to the egocentric origin between the observer's eyes. The object cannot be located until a distance along the line is allocated (in this example, 2 units).

Visual distance

Information about the direction of objects with respect to the perceiver is clearly important, but insufficient by itself to control anything but very simple sorts of behaviour, like guiding oneself towards a target. For example, the tendency (called tropism) of plants and animals to grow or move towards certain sources of stimulation, and away from others, can be based on a sensory estimate of the direction in which the stimulation is strongest. However, an object could lie anywhere along a line of visual direction, and for most behaviour it is necessary to know the distance as well as the direction of objects in the environment. Together, distance and direction specify the geocentric location of an object in three-dimensional space with respect to an observer. The size of an object cannot be known from direction alone, either. The lines of visual direction, which define the boundaries of an object, could correspond to any of a large number of identically shaped objects, with different sizes and distances. All that would be known is the angular size of the object, and this would not generally be sufficient for recognition or appropriate action.

For human behaviour, and probably that of most vertebrates also, perception involves recovering both the direction and the distance of objects in the environment, so that their location is specified in a geocentric frame of reference. Again, we encounter different terminologies that have been applied to similar perceptions. The perceived distance from the observer to an environmental feature has sometimes misleadingly been called absolute distance, although as we have seen this term can have no clear meaning, either physically or perceptually. All perception is relative to some frame of reference, none of which are absolute. The perception of distance from the observer to objects plays a very important role in the perception of size and shape, as will be discussed in Chapter 8. We shall refer to object distance for want of a more succinct alternative.

Relative distance is the term normally given to the perceived separation in space between objects. Although it is hardly a precise term since all distance

perception is relative to some frame of reference, it has been used so extensively that we shall also adopt it. If the objects are stationary, their relative separation remains constant, despite any movement by an observer. As will be seen, one of the critical requirements of the visual system is to recover relative distances and directions in the environment despite movements of the observer's eyes, head and body.

Just as with direction, there is always a frame of reference involved in any specification of distance, and it is important to make this explicit in discussing distance perception. We are familiar with physical measurements of geocentric distance, using devices like rulers and tape-measures, and scales with units like inches and feet, or centimetres and metres. These scales provide metric distance information, which conveys not only the order of objects in depth, but also the magnitude of the distances between them. One reason such scales are needed is that experience teaches us not to rely on judgements of distance 'by eye', if accuracy is required. Nevertheless, we obviously do perceive distances, whether accurately or not. If we take behaviour, based on perception, as the index then we are very accurate at such estimations for objects nearby. We can reach for objects and grasp them or avoid objects with which we might collide during locomotion. We are so good at these aspects of distance perception that we do not need to apply a cognitive value to the metric distance of objects. Despite this, most people can apply familiar units to express their perception of distances, and are able to report how far away something appears, for example in centimetres. No doubt this is due to our constant use of such scales since childhood. However the guide dog which featured in Chapter 1 would have no such familiarity, but could still perceive distances and act accordingly.

Essentially, a perceptual system should represent distances in a way that is lawfully related to physical distances in the outside world, at least for the range of distances which cover the space where we can carry out actions. In effect, this means that perception should conform to the nature of the three-dimensional space in which we live. Such a requirement is unlikely to be met with compete accuracy, or to hold for longer distances. However, it is hard to imagine behaviour of any complexity without effective distance perception. Many of the activities needed for the survival of humans and guide dogs would be impossible if perceived distances did not correspond to the physical space in which the activities are carried out. Since we have evolved to survive in a three-dimensional environment, it is hardly surprising that we have developed perceptual systems capable of representing that environment with sufficient accuracy to guide our behaviour.

It is striking that many of our units of distance have their origin in human anatomy, like the length of the forearm or foot, or the stride when walking. Possibly, in the absence of verbal scales, distance may be expressed in terms of the muscular effort needed to reach a given point in the environment. Even someone who had difficulty in stating the apparent distance of a target might be quite good at reaching out and touching it. Human distance perception is quite poor at longer ranges, such as more than a few metres, as is well known to drivers, pilots, mountaineers and golfers. Our ability to represent distances

decreases outside the range of action of an unaided individual, and this may have important consequences for our ability to function effectively, or even survive, in a mechanised environment quite unlike that in which our species evolved.

How do we come to perceive distance at all? This question has been at the heart of perceptual investigation for centuries, and it is generally answered by description of the various sources of information about distance which are available to us. These sources of information are known as cues, and their identification and analysis has been an important achievement, although primarily on the part of artists rather than psychologists. There are a number of these cues, which under normal circumstances give rise to a reliable perception of distance. When failures occur it is generally because of the lack of strong cues to distance, which causes inaccurate perceptions. The major cues to distance are described in the following two sections, in terms of their contribution to the perception of distance between the self and objects, and the relative distance between objects.

Object distance

Suppose that you are seated in a dark room, and a small illuminated disc appears in front of you. The disc will be at a particular metric distance from you, but at what distance, if any, will it be perceived? The apparent direction of the disc would vary, and it would appear to move around randomly (the autokinetic effect, see Chapter 7). However, its distance would appear fixed and it would not seem to move towards or away from you. Under such reduced conditions, your perception of distance may be quite inaccurate, but the disc will still appear to be at some distance, even if this is erroneous. In fact, in a situation like this, observers will see an object as being at a particular distance, generally around 1.5–2 metres, regardless of its physical distance. This has been called the specific distance tendency by the American psychologist Walt Gogel (1918–2006). The specific distance tendency ensures that some value of perceived distance will be present in the perception of any object. Generally, we are not strongly influenced by the specific distance tendency, because we rarely view objects in complete isolation. Gogel's work has demonstrated the strong influence of distance perception on the perception of other visual characteristics such as shape, orientation and motion.

One object distance cue, accommodation, was described in Chapter 4. In principle, the accommodative state of the eye could provide information about distance, although as was discussed earlier this will be confounded with the light level in the environment and with the wavelength of the light. The question of whether, and to what extent, accommodation actually does play a part in determining perceived distance remains controversial. Probably it is at best a weak cue, which comes into play mainly at short object distances, or when other sources of distance information are reduced.

Object distance can be determined from ocular convergence provided the distance separating the two eyes is known (see Figure 6.15). This interocular distance (IOD) averages around 65 mm in adults, but varies quite considerably between individuals. It also changes within individuals as the result of growth,

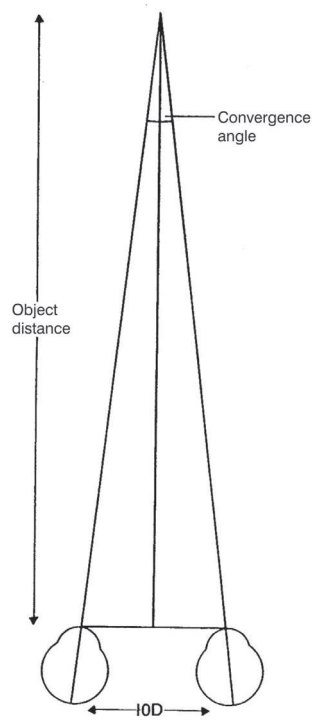


Figure 6.15 The geometry of ocular convergence and distance. The physical distance to the point on which the eyes are converged can be found from the ratio of the interocular distance (IOD) to the angle by which each eye deviates from the straight ahead position.

but presumably any changes are slow enough for us to adapt to them as they occur. Provided an object is straight ahead, its distance can be found from the ratio of the convergence angle to the IOD. More precisely, for this case, the distance is equal to half the IOD divided by the tangent of the angle by which each eye deviates from the straight ahead position. For objects away from the straight ahead position, the relationship is more complex, but still solvable. It is not likely that the nervous system actually performs the operation of taking the tangent of an angle, but rather that there is in practice a consistent relationship between the state of the muscles that move the eyes, and the distance of centrally fixated objects. Such a relationship would be established as the result of visual and manipulative experience with objects, and the nervous system is good at detecting and recording the way in which one sensory event is correlated with another. For convergence to be useful in anchoring relational distance it would only be necessary to apply it to an object that is visible, rather than assuming that convergence is applied to every object fixated. Once one value for object distance is derived from convergence it can be used to scale the other relational distances that are extracted.

There is some disagreement as to exactly how information about ocular convergence might be obtained; the alternatives are either feedback from detectors in the ocular muscles themselves, or more likely, a copy of the commands sent to the muscles from the brain. In either case, for convergence to be an effective cue to distance, such information must be available, from whatever source. In order to show that convergence affects perceived distance, it is necessary to devise experimental conditions in which convergence can be controlled and altered, and other cues are eliminated or held constant.

Changes in convergence can lead to paradoxical effects on the perception of distance and size, which have caused some controversy as to whether convergence is a reliable cue to object distance at all. Typically, if convergence is reduced so as to correspond to a greater distance, while the physical distance to the target remains constant, the perceived size of an object also increases. Similarly, perceived size decreases if convergence is increased. The paradoxical effect is that an observer may not also report that the perceived distance has changed, when the convergence does. Possibly we modify our judgement of distance on the basis of perceived size. A larger appearing object may be judged to be nearer, even though the information from convergence, which has determined the perceived size, indicates the opposite.

Vergence movements change the state of convergence to different object distances. Vergence movements are approximately equal in extent, but opposite in direction, in the two eyes. In addition, as discussed earlier, the eyes may move so as to change the point of fixation to a different egocentric direction. Such version movements cause the eyes to move through equal extents in the same direction. Both vergence and version obey Hering's law, which states that the eyes act as a single functional unit, whose movements are always of the same extent – in the opposite direction for vergence, and in the same direction for version. Hering's law applies independently to vergence and version, which seem to be controlled independently. This can be seen clearly if fixation is shifted from one target to another at a different distance and in a different direction, which requires a change in both vergence and version. There is a shorter delay before the vergence movements start than is the case for version, so there is an initial change in vergence towards the distance of the new target. However, vergence movements are relatively slow compared with version, and before the new state of convergence has been reached, a rapid change in version takes place, appropriate to the direction of the new target. The vergence movement then continues until the target is imaged on the fovea in each eye. In each phase, eye movements obey Hering's law, although the overall result may be that each eye has moved by a different amount in order to change fixation.

Information which is geometrically equivalent to convergence can also be obtained by viewing an object from different directions, for instance when the head is moved from side to side. The change in an object's egocentric direction when viewed from two positions could be used to find the object's distance, if the extent of the head movement is known. There is some evidence that this cue of egocentric motion parallax can be used by human perceivers.

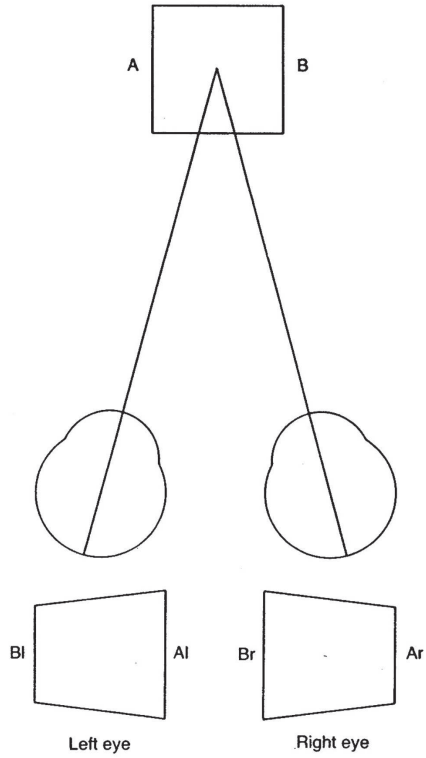


Figure 6.16 Vertical disparity produced by convergence. The stimulus is a square with equal sides, but its optical projection to each eye is unequal. The trapezoids represent the left and right retinal images of the square, and show that the temporal vertical sides are longer than the nasal ones. Binocular fusion takes place between Al and Ar, and between Bl and Br. The difference in the projected lengths of these pairs constitutes the vertical disparity. This provides a cue to the distance of the stimulus, since it varies with convergence.

It is also possible that we are sensitive to the differential vertical magnification of images in the two eyes. When an object is beyond about 6 m, it projects images whose vertical size on the retina is effectively the same for both eyes. However as an object is brought closer, the eyes must converge if the object is to be imaged on the foveas of both eyes. As a result there is a difference between the vertical extent of the retinal images in each eye. Imagine that the object is an outline square, perhaps drawn on a piece of paper. If the eyes are appreciably converged, the retinal images will be trapezoidal, with the vertical nearest the nose shorter than the vertical nearest the temples. As can be seen from Figure 6.16 however, the nasal and temporal lines represent opposite sides of the square in each eye. Thus the nasal vertical in the left eye corresponds to the temporal vertical in the right eye. This difference between the vertical extent of images falling on corresponding regions of each eye increases as an object is brought closer. This is a purely optical source of information about object distance, in that although it

arises as the result of convergence, its interpretation does not require information about the positions of the eyes.

One question raised by the above discussion is how we are able to see the distances of objects beyond around 6 m. Conventionally, rays of light emanating from a source at 6 m or more are said to lie at optical infinity, because they are effectively parallel to the optical axis. There will therefore be no appreciable convergence of the eyes when viewing an object at or beyond this distance. Clearly, the world does not appear to collapse into two dimensions beyond 6 m, so there must be additional processes at work. One possibility would be the familiar size of objects. If the things we see at a distance are familiar from past experience, such as people, buildings, or vehicles, then we might base the perception of their distance on our knowledge of their true physical size. This would require that the perceived size of the object is compared with some representation in memory of how big the object actually is. The ratio of perceived size to remembered physical size could then be used to derive the distance. The role of familiar size in distance perception has been the subject of many experiments. What seems to be the case is that the perceived size of objects, whether or not they are familiar, is based on their retinal size and their perceived distance. This gives a perceived size which may or may not correspond to familiar, remembered size. If there is a discrepancy between the size that a familiar object appears to be, and the size which we know it to have, then we may alter what we *say* we think its distance is. If you look down from a tall building towards people on the ground, there is very little distance information available. The people below will look small, despite our knowledge that they are not likely to be abnormal in this respect. Thus our perceptions will not be altered by our past knowledge, but our verbal interpretations and possibly our actions may be.

Relative distance

The various cues that indicate the relative distance between objects enable us to extend the perception of distance in the environment beyond the range of object cues, and also to obtain information about the three-dimensional structure of objects. In themselves these cues cannot give an object scale of distance or depth; they can indicate the relative separations of objects and parts of objects, but the scaling of the separation is ambiguous. Given three points at different distances, relative cues would enable an observer to state that the distance from A to B is twice that from B to C, but not to give metric information about these separations. If the points A, B and C lie on the surface of a solid object, relative cues will correctly specify its shape, but not its overall size. That is, it would be possible to recover the three-dimensional shape of a car, and recognise it as such, but not to say what size it was, or even whether it was a real car or a toy. However, if the object distance to any one point is known, perhaps from convergence or vertical disparity, then this allows relative distances to be scaled unambiguously. This means that relative cues can provide distance information extending out to a greater range than is possible with object cues alone. All that is needed is

the object distance to one nearby point, to act as a stepping-stone for the rest of the visual scene. In principle the relative distances between objects could be obtained from a series of object distances; if the distances to A and B are known independently, the distance from A to B could be given by subtraction. In practice, as we have seen, object distance information is not available much beyond the range of physical action, and its primary role seems to be to provide a metric for interpreting the cues that specify relative distances directly.

The first people to understand the importance of cues to relative distance were painters, rather than scientists. The problem is to establish how far such pictorial cues to distance are important for normal perception. Clearly, to the extent that they work in pictures, they are able to influence perception. What is more doubtful is whether we should regard the perception of pictures as representative of normal perception (see Chapter 8). This might be reasonable if the retinal image was a single static 'snapshot' of the external world, like a picture, but in fact it is not. Not only does perception derive, in those with two eyes, from the combination of two different images, but these images are in constant motion as the result of eye movements and head movements. Many textbook descriptions of distance and depth perception are primarily accounts of the perception of pictures, and this leaves out all the information that is available to an active, moving perceiver.

The general approach has been to determine the effectiveness of a relative cue to distance by presenting it in isolation, and assessing the extent to which it creates a perception of depth. In practice, experimental techniques cannot eliminate all information about depth other than the cue being studied. There is therefore a cue conflict, and the actual perception of distance is a compromise between the various sources of information, depending on their strength. A good example of this is the perception of depth in pictures. A picture may incorporate a number of cues to relative depth, such as overlap, height in the picture plane, perspective and shadowing. All these cues can indicate consistent depth, and need not conflict with each other. However they do conflict with other information that specifies a flat plane, arising from the picture surface, its frame and surroundings. The depth that we see therefore depends on a compromise between the various cues. Cues to the flatness of a picture can be reduced, for example by creating an image on a surface without a visible texture and without a clearcut frame. If so, the impression of depth in a picture can be markedly increased, as is the case with large screen projections in the cinema.

A compelling cue to relative depth arises when the images of an object fall upon different retinal areas in the left and right eyes (Figure 6.17). As described above, stimulation of corresponding retinal locations gives rise to single vision, and perception of a single visual direction. If the stimulated locations are very different, then double vision or rivalry results. Between these two extremes, there is a range where stimulation of non-corresponding points causes single vision and the appearance of depth. The process that gives rise to depth in this way is called stereopsis, and it is probably the most intensively investigated phenomenon in the history of vision research.

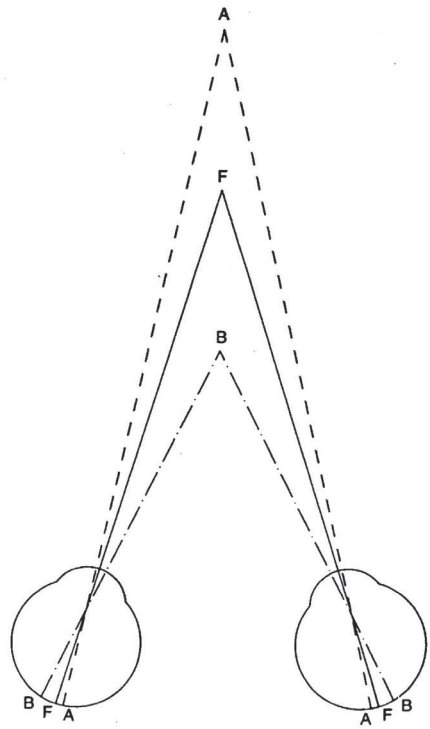


Figure 6.17 The geometry of retinal disparity. Both eyes fixate on point F, so that its image falls on the fovea of each eye. The images of points A and B, which are respectively further and nearer than F, fall on non-corresponding points. These differences in the projections from the two points are referred to as retinal disparities. If the disparities are not too great they can be processed in the visual system to provide information for relative distance, called stereoscopic depth perception or stereopsis. The sign of disparity depends on whether the non-fixated point is nearer or further than the fixation distance. Disparities can be described in terms of visual directions. Point A appears to the right of F when viewed by the right eye alone, and to the left with the left eye; this is called uncrossed disparity. Conversely, point B appears to the left of F when viewed by the right eye, and to the right with the left eye; this is referred to as crossed disparity. These differences in visual direction can easily be demonstrated. Hold both index fingers straight ahead of the egocentre, with one a few centimetres higher and further away from the other. If you fixate on the nearer one the far one will have uncrossed disparity, whereas fixation on the further one will result in crossed disparity for the nearer one.

The geometry of binocular disparity was described long before the function that it serves was appreciated. In fact, the link between disparity and stereopsis was not described until 1838, when the physicist Charles Wheatstone published his account of the stereoscope and the experiments he had conducted with it. A stereoscope is an instrument that permits presentation of slightly different figures to each eye independently. If the figures have the characteristics of the left and right eye views of a solid object, then the impression obtained is of a solid or three-dimensional figure rather than a flat one. There are several ways in which this can

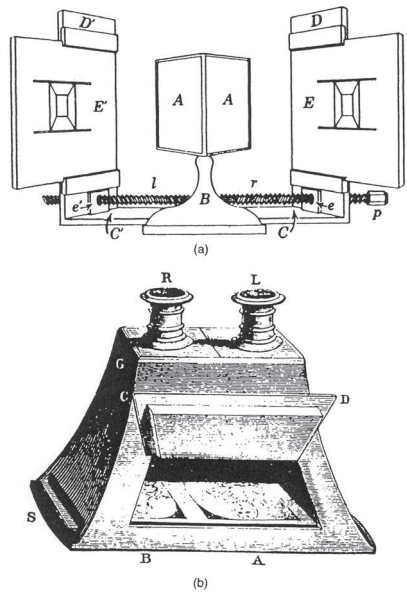


Figure 6.18 (a) Wheatstone's mirror stereoscope. The two slightly different perspective drawings of a solid object are placed on the side panels and their images are reflected to each eye via the mirrors (A). (b) Brewster's lenticular stereoscope. Half-lenses are located in the viewing tubes and act as magnifiers and prisms, so that the two pictures (A and B) appear to be straight ahead. This type of stereoscope was particularly popular in the latter half of the nineteenth century for combining stereophotographs.

be achieved optically. Wheatstone's stereoscope consisted of two plane mirrors, at right angles to one another, which reflect figures mounted appropriately on the side arms (see Figure 6.18a). An alternative model based on lenses was devised in 1849 by a Scottish physicist, David Brewster (see Figure 6.18b), and it was this type of stereoscope that graced many a Victorian parlour. The popularity of the stereoscope in the nineteenth century was based not only on the impression of depth derived from two flat pictures, but also because it could be married to the newly invented art of photography.

In addition to systems based on lenses or mirrors, there are other ways of viewing a stereogram in order to obtain stereopsis. The basic requirement is that a different pattern of stimulation can be presented to each eye independently. Some people can voluntarily alter their ocular convergence so as to fuse the two halves of a stereogram, and thus see the represented depth. This free fusion becomes easier with practice, and it can be tried with Figure 6.19. Anaglyphs are stereograms in which the left and right eye images are superimposed, but printed in different colours, such as red and green. When seen with a red filter over one eye and a green one over the other, the combined image is separated into a red and black image to one eye with a green and black image to the other, and binocular fusion may take place between the disparate black images. This method is also effective

Figure 6.19 Stereophotograph of a stereoscope. The method for combining images in different locations, described in the legend to Figure 6.9, can be used to fuse this stereopair. The left and right images have been arranged so that the one on the left should be viewed by the right eye, and the one on the right by the left eye. If these photographs were combined with a stereoscope, so that the left one went to the left eye, then the disparities would be reversed, as would the perceived depth.

with projected images, and has in the past been the basis of so-called 3-D films. While stereopsis in moving images is impressive, the anaglyph method is not entirely satisfactory because of the instability of binocular fusion with monocular images of different colours. The same result, but without the perceptual difficulties involved in colour separation, can be achieved with polarising filters and is employed in more recent stereo cinematography. The left and right eye images are seen by separate, simultaneous projection through polarising filters onto a single display screen; the filters are polarised in directions at right angles to one another. The images are viewed with a corresponding polarising filter in front of each eye and the binocular fusion is stable, although of reduced contrast. A more complex approach that achieves both stability and high contrast requires computer control. Each eye views the scene through a LCD filter that can be rapidly switched from transparent to opaque, in synchrony with rapid alternation of the left and right eye views projected on the display screen. When the left eye view is presented on the screen, the left filter is transparent and the right opaque, and vice versa. Alternation at 20 cycles per second or more gives a convincing impression of stable stereoscopic depth, but is both more expensive and less suitable for large audiences than anaglyphs or polarising filters. Yet another method is based on printing the two views in narrow, alternating vertical strips with small lenses in vertical strips over each one. The lenses are made so that the light from one vertical printed strip is refracted to reach one eye while its neighbour is seen by the other eye.

Wheatstone was able to free-fuse easily, but he made the mirror stereoscope so that those who had difficulties with free fusion could still experience stereoscopic depth from paired pictures. An even earlier description of depth perception without any optical device is called the wallpaper illusion, because regularly patterned wallpaper is an ideal stimulus for inducing it. If similar, but

laterally displaced, patterns are combined binocularly (by convergence at a depth plane in front of or behind the wall) then the surface is seen either in front of or behind the actual wall. This effect has been manipulated with great subtlety with autostereograms consisting of a multitude of small elements. Rather than having all the elements evenly spaced, autostereograms have some elements repeated at regular horizontal separations and others varying by precisely computed extents. When the repetitive elements are combined by over- or under-convergence then particular patterns emerge in depth. Like their close cousins, random dot stereograms, autostereograms are easy to see by some people but others require considerable time and practice to see the depth that they can elicit.

The geometry of binocular projections, as shown in Figure 6.17, was known and illustrated in the sixteenth and seventeenth centuries, but binocular vision was considered solely in terms of visual directions. An object that was further away than the point of binocular fixation appeared directed to the right side when using the right eye alone, and to the left side when using the left eye alone. This is now called uncrossed disparity. Conversely, a nearer object is said to have crossed disparity, because it appears directed to the left of the fixated point when using the right eye and to the right when using the left eye. Prior to Wheatstone, the principal problem in the study of binocular vision was that of singleness of vision: how is the world seen as single when we have two different views of it? The generally accepted answer was that only objects with the same visual directions would be seen as single: only objects stimulating corresponding points (or falling on the Vieth–Müller circle) would be seen as single, all others would be double. Wheatstone demonstrated experimentally with the stereoscope that this was false: the stimulation of slightly non-corresponding points can yield singleness of vision, but it also results in the perception of depth. Accordingly, Wheatstone established for the first time that binocular vision is concerned with distance as well as direction, and it is for this reason that the instrument he invented and the experiments he conducted with it are so important.

Binocular or retinal disparity is a relative cue, because it applies to situations where one object is fixated (and therefore stimulates corresponding retinal points) while one or more other objects at different distances project to different regions of each eye. As can be seen from Fig. 6.16, disparity can be produced either if an object is nearer than the plane of fixation, or further away. For an object of fixed size the disparity will vary with object distance. Wheatstone described the changes that occur when an object approaches the eyes (or when an observer approaches an object): there are increases in retinal size, retinal disparity, convergence and accommodation. Despite these changes the object appears to remain the same size. Wheatstone conducted experiments to manipulate these cues independently, and of particular interest was the relationship between disparity and convergence. When retinal disparity was kept constant and convergence increased, perceived depth decreased. Conversely, when convergence was held constant and disparity was increased perceived depth increased. Therefore, under normal circumstances there is a trade-off between increasing retinal disparity and increasing convergence which serves to maintain the constancy of perceived depth.

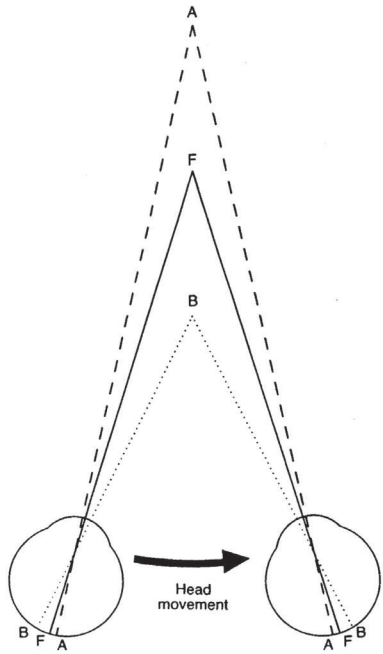


Figure 6.20 The geometry of relative motion parallax. The observer maintains fixation on point F, while moving the head from left to right through the distance shown by the arrow. The images of points A and B, which are respectively further and nearer than F, are displaced across the retina, by an amount which depends on the extent of the movement and the distance of the points relative to the plane of fixation. The distances AF and BF can be recovered if the object distance to F and the extent of the movement are known. The pattern of stimulation, and the necessary sources of information, are analogous to the binocular disparity produced by static observation with two eyes (compare Figure 6.17).

These points can be emphasised by means of an example. Suppose that an object such as a pencil is held at a slant so that the point is towards you, and that you are fixating on the point. There is a disparity between the retinal images of the other end of the pencil, which alters as the whole pencil is moved towards or away from you. The relationship follows an inverse square law; the disparity is proportional to the inverse of the object distance squared. Thus if the object distance to the point of the pencil is doubled, the disparity of the far end will be a quarter of its original value. Because of this, in order to perceive that an object is a constant length in depth, it is necessary to scale the disparity with the perceived distance. Errors in perceived distance will produce errors in the perceived depth from disparity. The IOD also determines disparity; for any given physical depth interval, the disparity will be larger if the eyes are further apart. As with object distance from convergence, the IOD must be available to the visual system.

While it is obviously necessary to view the world with two eyes in order to detect disparity, there is a functionally similar cue available to monocular vision,

if the head moves from side to side. This cue is relative motion parallax, and its geometrical basis can be seen from Figure 6.20. This figure is identical to Figure 6.17, except that the two eyes have been replaced by a head moving sideways through some distance. It is evident that the successive projections of objects on a single retina carry similar information about relative depth to binocular disparity. If you move your head to the right so that your left eye is located in exactly the position previously occupied by your right eye, then the successive retinal projections to your left eye are identical to those that were available to your left and right eyes, with your head in its original position. Experiments have shown that the depth produced in this way is very similar, if not identical to stereopsis, and it is possible that both patterns of stimulation activate the same cortical mechanisms. For example, prolonged viewing of a stereogram of a corrugated surface produces an aftereffect in which a flat surface appears to be corrugated in the opposite direction, and this also occurs with depth due to motion parallax.

Motion parallax consists of relative motion between the retinal images of objects at different distances, and is created by lateral motions of the observer's head. The kinetic depth effect (see Chapter 7) shows that relative motion may also act as a cue to depth even if seen by a stationary observer. Suppose an object is placed behind a back projection screen, and illuminated by a source of light such as a projector. If the object is stationary, its two-dimensional shadow projected on the screen will convey little or no information about the three-dimensional shape of the object. However, if the object is rotated, an observer will report the appearance of a solid object, rather than relative motions between parts of a flat shadow.

If one eye is covered, and the head is kept still, the visual world does not suddenly cease to appear three-dimensional. The cues available in this situation are primarily those which were referred to above as pictorial, because of their use in graphical representations. Despite the term, however, they are also present in natural scenes, which is why they are also effective in pictures. Several such cues have their origin in the fact that the retinal projection of any object or surface decreases in angular size with increasing distance. This fact of geometrical optics means that natural scenes contain perspective due to differences in relative retinal size between similar objects at different distances. Also, more distant objects on a ground plane are located higher in the visual field than nearer ones. Most natural surfaces have a visible texture, and man-made ones commonly have either decorative patterns or a regularly repeating structure. When seen at a slant such surfaces therefore display a gradual change in the density of the texture. Such texture gradients provide information not only about relative distance between the nearest and furthest parts, but also about the rate of change of distance; that is, edges, steps, corners and other surface features. Relative size may also change over time. If an object approaches you, its retinal projection increases in angular size. That this can act as a cue to relative distance is shown by the fact that observation of images undergoing expansion and contraction produces a compelling perception that the represented object is moving in depth, towards and away from the observer.

The constraints of optical geometry also ensure that opaque objects obscure more distant ones where they overlap, or modify the colour or brightness of the transmitted light if they are transparent. If the observer moves, different portions of obscured objects become visible, and this revealing or obscuring of texture can itself function as a cue to distance.

Over very long distances in the natural environment the colour of surfaces tends to become bluer. This effect, known as aerial perspective, is due to the absorption and scattering of shorter wavelengths by dust and moisture particles in the atmosphere, and gives a sense of receding panorama to a landscape. In very dry or clear air aerial perspective is reduced, and it may be hard to judge the distance and size of landscape features. Aircraft pilots and mariners report problems with navigation as a result. The importance of the various cues to distance in the natural environment is shown when they are reduced or eliminated; the perception of space then becomes very uncertain. In the absence of strong cues to distance, adjacent areas of the field of view tend to take on similar apparent distances, a phenomenon which is called the equidistance tendency. This effect can be quite pronounced if a distant scene is observed through a nearby frame, like a window. The strong cues available for the distance of the frame tend to make the whole scene appear to be at the same distance.

It is common for textbooks on perception to include an illustration showing the various cues to relative distance. We feel that this may be misleading, since the impression of depth in pictures is complex, and invariably based on a number of conflicting cues. Alternatively, the cues mentioned and illustrated above can all be seen by observing your own visual environment. If you look at a nearby object, such as a pen held at arm's length, you can note the difference between the views of the object available to each eye, by opening and closing each eye alternately. Using both eyes, converge on a surface about one metre beyond the pen, and you will see the double images of it, since they have too great a degree of crossed disparity to be fused. If you reduce the distance between the pen and the further surface, the double images will disappear and you should see a single pen and the background separated in depth. Now observe with one eye only, but move your head from side to side through about 10 cm. The depth should reappear, this time from motion parallax. With a stationary head and with one eye covered, look at the room you are in, or out of a window if possible. If there is enough distance, you can observe the presence of perspective in the texture of walls, floors or ceilings, and in the convergence of actually parallel contours. If there are a number of objects in view, some will partially obscure others, providing the cue of interposition. Note the presence of shadows; we generally tend to ignore their presence, but they carry information about the shape of the objects that cast them, or the surface on which they fall, if the location of the light source is known. Paradoxically, the fact that you can see these effects with monocular observation and a stationary head means that are they at best partially effective as cues to distance. If they were wholly effective, then they would cause the modification of perceived size to correspond to the actual physical distance, and no differences in relative size would be apparent.

Navigation

The issues discussed in this chapter have concerned determining our location in space – deriving a geocentric representation of the locations of objects in the surrounding environment relative to us. This provides the coordinates to control our behaviours like reaching, writing, and walking. Crossing the road engages all the processes discussed above as well as those for motion and recognition. However, these actions are often undertaken to accomplish tasks that can be beyond the scope of perception; that is, the goal or end point is not available to the senses. When we cross the road we can see the other side that we wish to reach, but when we go to catch a bus we may be far away from the route that the bus takes so that we need to navigate through the environment to get to the bus stop, detect the appropriate bus, and alight. We typically use mental maps of the environment to navigate successfully, and there is much debate about the form these mental maps take and how we use them. Are they comprised of landmarks that we link, do we have formal internal lists of directions that we follow, or do they conform to the layout of a geographical chart? How are objects and their spatial relationships represented? Charts or graphs have similar features because they represent arbitrary two-dimensional surfaces on which distances are compressed in some convenient way. For maps this is determined by the scale, for example 1 km on the earth's surface could be scaled down to 1 cm on a map. Locating a position on a map can be given by the values of latitude and longitude, both of which have to be derived from other measurements for those travelling by air or sea. It is somewhat easier to determine positions on graph paper because the positions of the axes (with respect to the earth's surface) are themselves arbitrary so that any point on the graph can be given vertical and horizontal coordinates. We tend to take maps and graphs for granted, although they are of relatively recent invention, and have afforded enormous advances in geography and geometry. Mercator devised the system of map projection that bears his name in the sixteenth century, and Cartesian coordinates (named after Descartes) were described in the following century. The maps used in antiquity often seem to have been maps of religious or political cognition rather than geography; nonetheless they apparently sufficed to guide political and military decisions.

It would seem clear that the characteristics of navigation might differ greatly from the geocentric perception that we achieve, but the concepts used to describe them should be similar. A distinction can be made between frames of reference based on the observer's position (which we have referred to as geocentric) and those adopting an origin outside the observer, sometimes termed allocentric or exocentric. Such an allocentric map would typically be two-dimensional and would include the position of the observer within it, so that navigation between various features of the environment could be achieved. One of the tasks for applied vision research is to determine the types of allocentric map (such as satellite navigation systems) that assist active observers in complex tasks like flying or driving; but these issues lie outside the scope of this book.

A navigating observer is, of course, moving through the environment. The objects in that environment may themselves change location in space and therefore also move. Objects and observers are located in time as well as space. Change in spatial location over time is motion, and the perception of motion is the subject of the next chapter.

Reference notes

John Locke's delightful example of how the location of chess pieces can be specified with respect to a series of frames of references can be found in his *Essay concerning human understanding* (1690, pp. 76–77). The translated quotation from Copernicus is from Copernicus (1978). Considerations of frames of reference have been commonplace amongst astronomers and physicists, but of surprisingly little concern to visual scientists. They are mentioned briefly by Kandel, Schwartz and Jessell (2000) and more extensively by Howard (1982) in the context of visual orientation. Orientation constancy is described by Day and Wade (1969). Unfortunately, there is no agreed terminology concerning frames of reference, and we might be accused of adding to the confusion! However, the system we suggest grew out of considerations of motion phenomena in general and the motion aftereffect in particular (see Swanston and Wade, 1992; Swanston, Wade and Day, 1987; Wade and Swanston, 1987, 1996). The one frame of reference term that is consistently applied to vision is 'retinotopic' or what we have referred to as retinocentric. Lotze's local signs were essentially retinocentric coordinates (see Rose, 1999). There is much less consensus concerning the frame of reference operating after retinocentric mapping. In the context of visual direction, Howard (1982) used the terms oculocentric, headcentric and bodycentric; these are similar to the terms we have used but they are not identical. Orientation seems an ideal aspect of vision to distinguish between the retinocentric, egocentric and geocentric frames of reference. It also provides a platform for introducing the concept of a patterncentric frame of reference (see Wade, 1992b). Our own research has demonstrated the operation of processes at this level in the motion aftereffect (Swanston, 1994; Swanston and Wade, 1992; Wade, Spillmann and Swanston, 1996).

Human visual orientation is the title of the excellent book by Ian Howard (1982), and he has discussed the factors affecting our perception of this dimension in chapters and articles (see Howard *et al.*, 2005). He has examined spatial orientation in a tilted environment like that shown in Figure 6.6. The tilted room in the photograph is located in Lingelbach's Scheune, in Germany; we thank Bernd Lingelbach for the opportunity of experiencing this disorientation. The tilted room presents a mismatch between the direction of gravity signalled visually and by the vestibular system. Much of the research on visual orientation has been stimulated by the demands of space flight (see Clément, 2011), where the gravitational reference is no longer operating, and all orientations are defined in egocentric or patterncentric terms.

As soon as binocularity is taken into account then the inadequacy of retinocentric mapping alone in determining visual direction becomes evident. This

was appreciated by Wells (1792) and Hering (1868/1942); their research has been integrated and presented by Ono (1981, 1990; Ono and Wade, 2012) and Howard (2012). The concept of the cyclopean eye was introduced by Hering. Howard's book provides a historical introduction to binocular vision, as do Wade and Ono (2012). Both examine Porta's (1593) suppression theory and the alternative fusion theory proposed by Aguilonius (1613) and the developments that have followed this distinction. Porta suggested a test of eye dominance and the development of ideas surrounding this concept are given in Wade (1998b). Until Wheatstone (1838) described stereoscopic depth perception the principal concern had been with singleness of vision with two eyes. Panum (1858) sought to reconcile the two views by proposing that the Vieth–Müller circle was not as restricted as is shown in representations like Figure 6.14; rather than a circle there is a region either side of the circumference within which disparate objects can be seen as single. Panum also introduced one of the stimuli most commonly used to study binocular rivalry – orthogonal gratings presented to the eyes, rather like those in Figure 6.13. He noted that rivalry is not always between the eyes, but parts of the patterns can be seen in each eye. Binocular rivalry continues to fascinate students of vision because variable percepts are a consequence of constant stimulation. Phenomena and theories associated with rivalry are discussed in Alais and Blake (2005). Brewster (1856) provided a counterpoint to Wheatstone, and all their papers on vision were published by Wade (1983). In addition to the horizontal disparities described by Wheatstone, vertical disparities can also play a role in depth perception (see Matthews, *et al.*, 2003). The contemporary concern with stereoscopic film and television has focused interest on issues of binocular vision and the techniques that can be used to produce it are described by Blundell (2011).

The topic of visual distance lies at the heart of art and philosophy. Artists have attempted to simulate distance on a flat surface for thousands of years, and many of the monocular sources of relative distance are called 'painters' cues'. The importance of binocularity was stressed by Berkeley (1709), although he enlisted convergence rather than retinal disparity to yield depth. The books by Sekuler and Blake (2005) and Coren, Ward and Enns (2004) describe the depth cues in distance perception. Both Wheatstone and Helmholtz recognised that motion parallax can operate as a source of information about relative distance, and this has been experimentally confirmed by Rogers and Graham (see Rogers, 2009). The history of the relationship between motion parallax and depth perception is outlined in Ono and Wade (2005).

In the context of navigation, Howard (1993) has made a distinction between frames of reference based on the observer's position (what we have referred to as geocentric) and those adopting an origin outside the observer (which he calls exocentric). Navigation is often considered as part of spatial cognition and as such is beyond the scope of this book. Nonetheless, writers on spatial cognition necessarily address issues of frames of reference (see Burgess, 2006).

7 Spatial motion

It might be thought that motion perception is simply the outcome of retinal image motion, but in fact this is so only for a stationary observer with unmoving eyes. Retinal image motion can be a consequence of the displacement of an object relative to a stationary environment, of an observer relative to a stationary object or of some combination of both. All these conditions result in changes in the pattern of stimulation at the eyes, but despite these we retain an appropriate representation both of our position in space and that of the object. Occasionally, errors do occur in our perception of motion in the natural environment, and these can be very instructive in understanding the nature of motion perception. It is these errors that have received most attention both historically and in contemporary research, which is perhaps surprising as this focus has diverted attention from the extraordinary ability of perceptual systems to recover motion accurately under widely varying conditions. We, as well as our metaphorical guide dog, can perceive the constant properties of objects (such as their location, motion, size and colour) despite large variations in the pattern of sensory stimulation.

Such errors in motion perception were remarked upon in antiquity before the basic perceptual processes were either described or appreciated. This was so because it was possible to compare observations of the same objects over time and to note any discrepancies between them. The modern definition of illusions refers to differences between the perception of objects and their physical characteristics. Consensus concerning an external reality did not exist in antiquity, and so attention was directed to those instances in which changes in perception occurred. That is, when the same object appeared to have different properties under different conditions. According to this observational definition of motion illusions, all that is required is an assumption of object permanence; thereafter, any changes in the appearance of the same object will be classified as illusions.

Motion phenomena

Many of the observational accounts of motion perception in the past involved such instances of discrepancies between objects seen under different conditions. Similarly, ideas about thresholds for detecting motion were also based on observations, but over longer periods. The slow motion of celestial bodies, like

the moon, were evident when observations separated by minutes or hours were compared, but not when the moon was observed directly. Since rotation of the earth was also too slow to be detected by the vestibular system, all the visual displacements were allocated to the moon rather than to the earth. Some celestial bodies did appear to move, sometimes very rapidly. These were shooting stars and their visible tracks across the sky were attributed to the slowness of vision, or what came to be called visual persistence.

John Locke, whose reflections on the spatial locations of chess pieces set the scene for the previous chapter, also made cogent remarks about spatial motion. He considered the limits of motion perception involved in the situations mentioned above – not perceiving motion of the moon and seeing the rapid motion of shooting stars as a streak. Locke wrote: "Motion can neither be, nor be conceived without space … *Motions very slow*, though they are constant, *are not perceived* by us; because in their remove from one sensible part towards another, their change of distance is so slow, that it causes no new *Ideas* in us; but a good while after one another: And so not causing a constant train of new *Ideas*, to follow one another immediately in our Minds, we have no Perception of Motion … On the contrary, *things that move* so swift, as not to affect the Senses distinctly with several distinguishable distances of their Motion, and so cause not any train of *Ideas* in the Mind, *are not* also *perceived* to move".

We will examine a range of motion phenomena by introducing them in a historical context and then addressing their contemporary importance. Movements that we can detect, like a car travelling along a road, often elicit eye movements in their pursuit. Eye movements have often been proposed as the basis of motion illusions. They will be discussed in more detail later in the chapter.

Induced motion

Induced motion refers to the apparent visual movement of a stationary object, usually in the opposite direction to the real movement of other objects. It can be produced with a wide range of stimulus configurations (Figure 7.1), and it occurs with linear and rotary motions in the frontal plane and with motion in depth. Induced motion is a contrast effect in the motion domain: if several objects move together and another remains stationary, then the latter will appear to move in the opposite direction.

A clear description of a naturally occurring instance of induced motion was given by Lucretius (c. 98–55 BC): "when the winds carry scattered clouds across the sky in the night time, then the shining stars seem to glide against the clouds and to pass above them in a very different direction from their true one". A distinction is implied between the real and apparent motion of the stars, presumably because the stars appeared to remain stationary when the clouds were not visible. Similarly, Ptolemy in the second century noted the phenomenon with regard to water flowing past a stationary boat: "When a boat is stationary on a calm river flowing without waves, a passenger who does not look at the distant land (but only at the water) thinks that the boat is moving quickly and that the water is still … Since the water

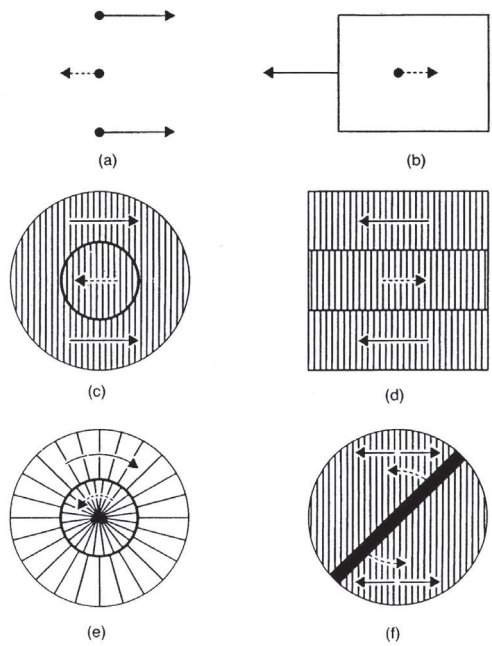


Figure 7.1 Laboratory stimuli that are used to induce movement in a stationary target. In all cases the physically moving parts are shown by solid lines and the dashed lines indicate the induced motion of the stationary components. (a) Two moving dots, (b) dot and rectangle, (c) annular movement of a grating, (d) flanking movement of gratings, (e) rotary motion of a radial annulus, and (f) line rotation induced by optical expansion of a grating. Linear induced motion is produced by (a)–(d), rotary induced motion by (e), and a combination of linear and rotary induced motion in (f).

appears stationary, the motion must be seen as coming from the boat. However, if the water, land, and boat are observed at the same time, and if we are aware that the land is stationary, we will see the water in motion". His account displays an appreciation of the relativity of motion perception with respect to the frames of reference that are in operation.

This observation has been repeated and reported many times in the last millennium, as it provides one of the most readily visible instances of induced motion in the natural environment. For example, William Porterfield described the natural occurring instance and gave a generalisation from it: "If two or more Objects move with the same Velocity, and a third remain at rest, the Moveables will appear fixed, and the Quiescent in Motion the contrary Way. Thus, Clouds moving very swiftly, their Parts seem to preserve their Situation, and the Moon to move the contrary Way".

Experimental studies of induced motion commenced rather late, considering this rich history. Motion can be induced with a simple arrangement consisting of a pair of moving and a single stationary stimulus (usually dots), as shown in Figure 7.1a. When only the dots, and nothing else, are visible then the physically

stationary dot is seen as moving in the opposite direction. A more compelling stimulus for inducing motion in a stationary dot is to displace a surrounding frame (Figure 7.1b). Under these circumstances the rectangle can be moved imperceptibly (i.e. below its motion detection threshold) and still induce motion in the dot. The possibility that the inducing motion resulted in unconscious eye movements was still entertained, and both psychophysical and physiological tests of this hypothesis were examined. In addition, rotation can be induced in a stationary radial pattern by rotation of a patterned annulus or surround (Figure 7.1e). The eye does undergo a rotation (torsion) as a consequence of surround rotation, but it is of limited extent and so is unlikely to be the basis of induced rotation. Evidence from measures of eye torsion confirms this conclusion – stimuli that modulate the degree of torsion that can be measured do not influence the magnitude of induced rotation reported.

The initial interpretations of motion phenomena are often in terms of eye movements. It reflects a desire to link an objective expression of motion (of the eyes) with a subjective experience of motion. If the two are correlated then it is assumed that the perception has been accounted for. That is, if motion over the retina (as a consequence of eye movements) is equated with perceived motion then the former is taken to account for the latter. Motion with respect to the eye (or the retina) is taken as the most basic interpretation of motion perception, and it does not require any understanding of subsequent processes (like those in the visual pathways or brain). It reflects a concern with aspects of perceptual experience (motion perception) and associated changes that can also be experienced subjectively or observed in others (eye movements).

Another psychophysical technique involves inducing motion in opposite directions simultaneously. For example, with fixation on a stationary central point, other stationary points to the left and right of it can appear to move outwards if the rectangles surrounding them move inwards. It is difficult to argue that an eye can move in opposite directions simultaneously and so this observation casts doubt on the link between eye movements and induced motion. Finally, direct measures of eye movements and linear induced motion fail to demonstrate their dependence.

A more subtle version of the eye movement hypothesis is that it is either uncompensated eye movement signals or tendencies for the eyes to move that are correlated with the induced visual movement. These have the same problems in dealing with apparent rotations, motions in depth, and opposite linear motions as do simpler eye movement interpretations. Accordingly, the perception of motion in a stationary stimulus induced by a larger moving one is unlikely to be explained by any correlated movements of the eyes.

Another type of motion interaction, between a moving field and its fixed surround, is called the barber-pole illusion. The sign for a barber's shop is traditionally a pole with a white and red helix painted on it; when it rotates the stripes appear to be moving along the pole rather than rotating. This is because we cannot detect the motion within the white or red areas, but we can detect the motion of the contours defining their boundary. Suppose a field of random dots moves horizontally within a rectangular frame. The direction of motion will be

seen unambiguously, and will not be altered if the frame has different orientations or shapes. There is a different outcome if the random dots are replaced by a pattern of black and white stripes oriented at an angle to the surrounding frame (hence the term barber-pole illusion, although in a real barber-pole the stripes are usually red and white). Now, the direction of motion is usually seen as aligned with the longer side of the frame, no matter what the physical direction of motion of the stripes may be.

Motion aftereffect

Induced motion is an example of simultaneous motion contrast, whereas the motion aftereffect reflects successive motion contrast. Having observed a moving stimulus for some time, a stationary one appears to move in the opposite direction. The motion aftereffect has a less even history than that for induced motion, but similar interpretations have been applied to it. Aristotle's descriptions of the phenomenon were mentioned in Chapter 2 but he did not relate it specifically to eye movements, nor was his description unambiguous regarding the direction of motion. The description given by Lucretius was more precise, and implicated the motion aftereffect with induced motion: "When our spirited horse has stuck fast in the middle of a river, and we have looked down upon the swift waters of the stream, while the horse stands there a force seems to carry his body sideways and pushing it violently against the stream, and, wherever we turn our eyes, all seems to be rushing and flowing in the same way as we are". Thus, when he was surrounded by flowing water and looking down at it the horse seemed to be moving against the current. This is an instance of induced motion. When he looked at the bank of the river it seemed to be moving in the same direction as the induced motion – that is, opposite to the flowing water.

Stimuli that can be used to generate motion aftereffects in the laboratory are shown in Figure 7.2. Unlike induced motion, the motion aftereffect seems to have drowned in the waters of antiquity. Indeed, it did not resurface until the nineteenth century, when Purkinje's acute vision was directed to it, although some earlier observers came very close to describing it. Purkinje described the motion aftereffect that follows observation of a variety of motions such as rotation and translation, and he related these to eye movements: "Another form of vertigo can be demonstrated if one observes a passing sequence of spatially distinct objects for a long time, e.g. a long parade of cavalry, overlapping waves, the spokes of a wheel that is not rotating too fast. When the actual movement of the objects stops there is a similar apparent motion in the opposite direction". Purkinje's interpretation was elaborated in 1825: eye movements generated by moving objects continued when stationary ones were subsequently observed. An unconscious eye movement explanation is explicitly adopted: the eyes pursue the moving objects and continue to do so with respect to stationary ones observed later. Addams adopted a similar interpretation. Both Purkinje and Addams describe what would now be called optokinetic nystagmus. The description by Addams is much more precise than that by Purkinje, as it refers to the cyclical nature of

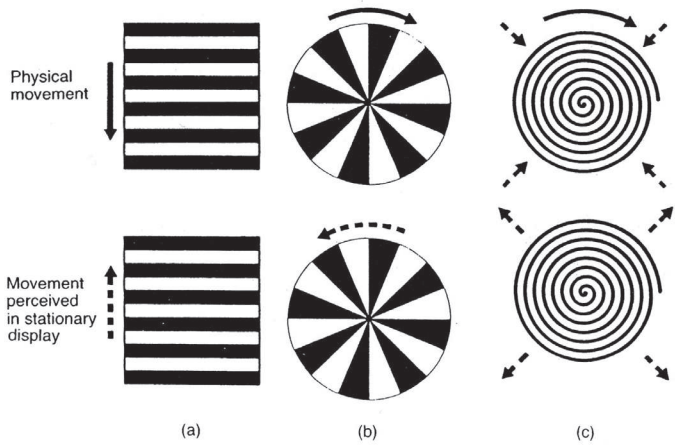


Figure 7.2 Laboratory stimuli that are used to produce movement aftereffects. (a) Linear motion of a grating, (b) rotation of a sectored disc, and (c) rotation of a spiral. The physically stationary patterns are initially seen as stationary. During the physical motion the movement is perceived as shown by the solid arrows, but the spiral also appears to contract. Following about 30 seconds adaptation the stationary patterns appear to move in the directions indicated by the dashed lines.

the movements and the frequency with which they occur. Moreover, Addams is explicit in relating involuntary eye movements and retinal displacements, which, he suggests, are the basis for the phenomenon.

The involvement of eye movements could have been excluded at an early stage, because of the stimuli that can produce it, but the conclusion was not made. Plateau (see Figure 2.26) introduced the stimulus employed most widely throughout the second half of the nineteenth century – a black disc with a white Archimedes spiral on it; it became called the Plateau spiral (like Figure 7.2c). Plateau described the phenomenon in the following way: "If the disc rotates in the direction indicated by the arrow, and one looks at it with the eyes fixed on the centre for a time sufficiently long, but not long enough to tire the eyes, then one immediately directs the eyes to another object, such as the face of a person, for example, one experiences a singular effect: the head of the person appears to shrink for some time. If the disc is turned in the opposite direction, the resulting effect is opposite: it is as if the head of the person appears to expand". That is, the visual motion is radial, either centrifugal or centripetal, which would be difficult to match with any eye movements. As was remarked: "The spiral of Plateau sounded the death knell for the eye-movement theory", but the tolling of this bell was clearly not heard by many.

Despite his awareness of Plateau's work, Helmholtz also adopted an eye movement interpretation of the motion aftereffect, although he did not describe formal experiments on the phenomenon. There is a logical difficulty of accounting for the phenomenon in this way: any after-movement of the eyes would affect all contours in the visual field, rather than being restricted to regions exposed to prior

motion. Motion aftereffects are generally restricted to the regions of the retina that have been exposed to motion. Mach pointed to the problems associated with eye movement interpretations and the difficulty they would have in accounting for radial motion seen in the spiral aftereffect. One of Mach's students had shown that counter-rotating spirals produce aftereffects that are centrifugal and centripetal at the same time. Moreover, rotating a central sectored disc in the opposite direction to two adjacent ones results in oppositely directed motion aftereffects that are visible simultaneously. Similar conclusions were reached from linear motion aftereffects. When horizontal and vertical gratings are moved vertically and horizontally, respectively, behind a circular aperture then the motion seen during adaptation is in a diagonal direction, with a motion aftereffect in the opposite diagonal. Motion aftereffects were even induced in opposite directions in the two eyes; when both eyes viewed the stationary test pattern the aftereffects engaged in rivalry, but each motion aftereffect could be seen when the other eye was closed.

Eye movements can be discounted experimentally because the motion aftereffect occurs with stabilised retinal images that remain in a fixed position on the retina and in addition eye rotations are not correlated with the rotary aftereffect.

Autokinetic sensations

Yet another motion phenomenon relates to the apparent movement of stars at night. In 1799 the great German naturalist, Alexander von Humboldt (1769–1859), recorded that when he observed certain stars they appeared to wander about haphazardly in the night sky: "Here is the place, at least in passing, to mention another optical phenomenon which I observed only once in all my mountain climbing, which occurred before sunrise on 22 June 1799 on the slopes of the Pics of Tenerife. In the Malpays, at a height of 10,700 feet above sea level, I saw with my naked eye lowlying stars in a wonderful wandering motion. Illuminated points rose upwards, moved sideways and fell back to their original positions. The phenomenon lasted 7 to 8 minutes and came to a halt long before the appearance of the sun above the horizon. The same motion could be seen through a telescope; there remains no doubt that it was the stars themselves that moved". The observation was made during a scientific voyage to South America, and it is clear that Humboldt genuinely believed that the stars did undergo the rapid motions that he perceived. He even called the phenomenon "Sternschwanken" or star wandering. Other observers were able to see the phenomenon, too, though they did not all describe the haphazard movements in a similar way. A debate arose about the precise basis for the star movements, and its subjective nature was determined by a group of astronomers who observed the same star at the same time and noted the directions of motion: all saw motion but not in the same directions. As there was no consistency in the motions of the same star seen at the same time, it became clear that it was a perceptual rather than a physical phenomenon. The phenomenon is not confined to stars in the sky, and the effect can be simulated with artificial stars in a darkened room. If you look at a stationary and very dim light in an otherwise totally dark room it will appear to wander about haphazardly,

like the stars that Humboldt observed. It is now called autokinetic motion. Most early reports noted that there is a latency of some seconds before the autokinetic motion is seen, and thereafter the amplitudes can be considerable. There was, however, much debate about the pattern and velocity of the motion seen.

Aubert rediscovered autokinetic motion in 1887. Some years earlier, Aubert had imposed some degree of order on a wide range of visual phenomena, particularly those involving motion. For example, he measured the thresholds for detecting uniform and relational motions. When a target moved in isolation its motion could be detected when it reached a velocity of about 0.3 deg/s. When a patterned background was visible the threshold was approximately 0.03 deg/s. Since isolated stimuli have higher thresholds for motion they can only be seen to move when they undergo quite large displacements. Thus, there is clearly an influence of the visible reference frame in motion perception, but that was not considered to be the sole factor. Aubert returned to the examination of motion phenomena in the 1880s, and rediscovered the apparent motion of a static light in a dark room; he gave it the name autokinetic sensation.

As with the phenomena mentioned above, eye movement explanations were the first candidates for the autokinetic sensations, but experimental investigations failed to find a simple link. One technique applied to examine this is to compare the movements of the afterimage (which is stabilised on the retina) with those of a spot of light. Under these conditions the afterimage remained relatively stable, even during quite large apparent movements of the spot. Using direct recordings of eye movements, their direction and amplitude did not correlate with those of autokinetic motion. However, there were some aspects of the phenomenon that could be related to the direction of eye movements, for instance autokinetic motion occurring in the same direction as prior eccentric fixation.

This intriguing phenomenon requires a dim and isolated light source in otherwise darkness. That is, there is no patterncentric stimulation relative to which the light source can be anchored. In like manner, when motion thresholds are measured for isolated lights their values are much higher (therefore sensitivity is lower) than when a visible background is present.

Stroboscopic motion

The aspects of motion perception discussed above are instances of the apparent motion of physically stationary objects. The converse applies to visual persistence: a rapidly moving object can appear stationary if its path is repeated. Under normal circumstances the moving luminous objects we observe can be resolved during their motion, either because they move slowly or because they can be pursued by the eyes. Certain naturally occurring events (like comets and lightning) do not satisfy these conditions, and they were commented on in antiquity. Similar effects can be seen with a rapidly rotating flame: it is seen as a circle. This phenomenon has long been known and it is a theme that was replayed constantly over the centuries: rapid motion of a burning stick results in the visibility of the path through which it passes. Leonardo drew circles of flame in this way; he

also added the subtle variation of moving the eyes with respect to a fixed flame, with a similar visual effect. Newton appreciated that the phenomenon could be employed to measure the duration of the persisting images. In the second edition of his *Opticks*, he suggested that its value was less than a second, but precise measurements were made in the eighteenth century by Chevalier Patrice D'Arcy (1725–1779). He constructed a machine with rotating arms onto which a glowing coal could be attached. The device must have been very large, as it was viewed in a large, dark room at a distance of about 55 metres. By measuring the velocity required to complete a visible circle of light (in an otherwise dark room), he calculated the duration of visual persistence to be about 130 ms – a value that corresponds closely to those subsequently obtained with more modern devices.

At the beginning of the nineteenth century a bewildering variety of such instruments were produced in order to exploit visual persistence. There were three ways in which this was achieved. Firstly, by rendering changing patterns of stimulation visible during continuous viewing, as in D'Arcy's rotating ember. Second, by brief presentation of parts of a single figure. Third, by successive stimulation with slightly different figures. It is the last that had the most impact on visual science. In the second case, an aperture moved in front of a scene or picture. A variant of this was the thaumatrope or wonder turner, a very simple arrangement invented by John Ayrton Paris (1785–1856) in 1825: a circular piece of card had different drawings on each side, and its ends were connected by string. When it was whirled both designs were seen superimposed: rats could be seen caged during rotation but free when the thaumatrope was stationary; fragments of words written on each side of the disc could be rendered complete due to visibility of their persisting parts. Visual and verbal puns were intentionally incorporated into the two elements.

The third mode of exploiting visual persistence was suggested by Michael Faraday (1791–1867): "The eye has the power, as is well known, of retaining visual impressions for a sensible period of time; and in this way, recurring actions, made sufficiently near to each other, are perceptibly connected, and made to appear as a continuous impression". This speculation proved particularly productive as it led to devices that would synthesise visual motion, one of the first of which is shown in Figure 7.3. Stroboscopic motion refers to perceiving motion from a rapid sequence of spatially separated static stimuli. It differs from real motion, in which objects change their location over time. It was with the invention of instruments like the stroboscopic disc in 1833 that motion perception was rendered experimentally tractable. Since the 1830s, stroboscopic motion has been one of the most intensely studied areas of motion perception. It has special interest, not just because of the significant practical applications, but also because it is an entirely artificial mode of perception. Unlike depth perception with stereoscopes (invented at about the same time), stroboscopic motion perception has no analogue in the natural world.

The apparent motion seen with the phenakistoscope or stroboscopic disc, and in films, is based on two phenomena – visual persistence and stroboscopic motion. Much of our artificial lighting is intermittent rather than continuous, but we are

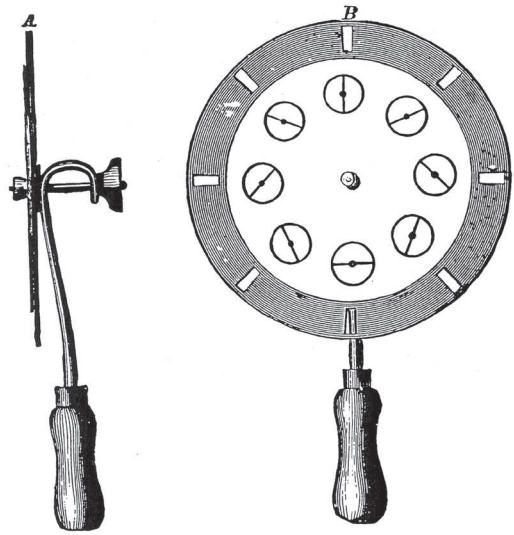


Figure 7.3 A phenakistoscope or stroboscopic disc. A side view is shown on the left; the disc could be rotated by turning the central knob. The front face is shown on the right. The disc is rotated while facing a mirror. Its reflection is observed through the slits around the circumference. Each slit is opposite a slightly different drawing – here a set of diameters – which are seen in succession, creating an impression of a rotating diameter.

generally unaware of this because of visual persistence; fluorescent tubes flicker with twice the frequency of the alternating current supply, but we only notice the intermittency when they become faulty. The frequency of flicker required to give a continuous (fused) impression can vary with environmental conditions and different individuals, but perception is normally continuous at or above 50 Hz. In the nineteenth century stroboscopic motion was a novel phenomenon: when two successive stimuli are presented in slightly different locations motion is seen, within certain temporal limits. If the blank interval between the briefly presented stimuli is too short then they appear to be visible simultaneously; if it is too long they appear to be successive, and at around 100–200 ms motion appears. It is possible to satisfy the conditions for stroboscopic motion independently of those for flicker fusion. This is what happened in the early days of the cinema. Films were taken and projected at 16 frames per second; this satisfies the criteria for stroboscopic motion, but it is rather slow for visual persistence, and so the pictures appeared to flicker – hence 'the flicks'. The problem was overcome by both increasing the frame rate to 24/s, and illuminating each frame with three pulses of light, so that the flicker rate is 72/s, which is well above our flicker threshold.

In the early twentieth century, Wertheimer built upon these phenomena of motion seen following successive stimulation of two spatially distinct stimuli, and he gave them a new interpretation. When the interval between the two

stimuli was around 100 ms the motion was optimal. He argued that the perceived motion was a Gestalt that could not be divided into smaller units. In furthering this theory, he examined the possibility that eye movements could be involved in stroboscopic motion. Wertheimer presented several arguments against the implication of eye movements. First, the duration of tachistoscopic exposure was shorter than that required to initiate an eye movement. Second, a fixation point remained stable during apparent motion; this applied to an afterimage, too. The strongest evidence derived from an experiment in which apparent motion was produced simultaneously in opposite directions. Wertheimer also demonstrated that there was no perceived distinction between apparent motion between two spatial locations and real motion.

Stroboscopic motion, like induced motion, can occur in different directions at the same time. This renders any interpretation in terms of eye movements difficult to sustain. Stroboscopic motion can also occur simultaneously from a central point in radial directions, and so the arguments against the involvement of eye movements that applied to the motion aftereffect are pertinent here, too.

Kinetic depth effects

There are various situations in which motion may be perceived as depth and structure in the moving object. If a wire frame shape is set up behind a screen and illuminated so that its shadow was seen by the observers, its shadow appeared as a two-dimensional pattern of lines when the object was stationary. When the object was rotated, the shadow was seen as a solid three-dimensional shape in motion. There are many possible interpretations of the pattern of motion of such a shadow, but the one most readily seen corresponds to the rotation of a rigid object. There are many examples of this process of the recovery of structure from motion and it is remarkable how little information is needed to produce perception of a stable three-dimensional shape.

Figure 7.4 illustrates patterns that may be used to create another type of depth perception from motion. If patterns like these are spun about their centres, a shape like a cone may be seen to emerge. Usually the apex appears to point to the observer, but it may also seem to be concave and receding. The apex of the cone appears to ‘wobble’ along a circular path. The effect is striking, and the

Figure 7.4 Patterns which, when rotated, produce strong kinetic depth effects.

kinetic depth effect was examined experimentally in visual science and it was applied with great skill to art by Marcel Duchamp (1887–1968). He called his productions rotoreliefs, and they were made between 1925 and 1935, when he published twelve rotary designs. Duchamp maintained an active interest in optics and motion, and the rotoreliefs provided him with evidence of the ways in which vision can be confounded.

The phenomenon of motion parallax, described in the previous chapter, provides a further example of perceived depth from motion. In this case the depth is most clearly and reliably perceived when the relative motion in a display is contingent on lateral movements of an observer's head. However, depth may also be seen by a stationary observer, in a similar manner to that obtained with the shadow-caster described above.

Visual vertigo

The early nineteenth century studies of motion aftereffects were stimulated by the investigations of visual vertigo, although the dizziness and giddiness caused by alcohol, looking down from heights, as well as a variety of ailments were known in antiquity. Reports of visual vertigo, like those of induced motion, have a more or less continuous history since antiquity. This probably reflects the fact that the experience is a common accompaniment of disease, and so has retained an interest for medical reasons. Vertigo can easily be induced in healthy people by rotating the body rapidly and then stopping suddenly: the whole visual world appears to be rotating in the opposite direction to previous body rotation.

Research on postrotational visual motion commenced in earnest in the eighteenth century, and it is a motion phenomenon that can be linked to eye movements.

The modern era was heralded by a much neglected Scottish/American scientist, William Charles Wells (1757–1817). He conducted many experiments on body rotation and described his results as follows: "When we stop ourselves while giddy from turning, our eyes do not return to the state of rest along with our bodies, but continue to move for some time after. Of this, however, we are not conscious; and hence we imagine the relative motion between our eyes and objects at rest to be possessed by the latter". The elegance of Wells' experiments lay not only in the description of the phenomenon but also in the use of afterimages, generated before rotation, to render visible the motions of the eyes after rotation ceased. Wells can be considered to have initiated the modern study of eye movements by seeking a more precise method of measuring them, namely by using afterimages as stabilised retinal images. He described the phases of nystagmus and the suppression of vision during the fast phase; he related the direction of eye movements to that of prior rotation; and he demonstrated torsional nystagmus. Moreover, Wells explicitly related the movements of the eyes to those of the visual world following body rotation. As would be expected of a phenomenon that is related to eye movements, visual vertigo involves a rotation of the whole visual world, rather than a part of it.

Vection

The rail system was also developing rapidly in the early nineteenth century, and rail travel exposed people to patterns of visual motion that were outside the natural range. It is not surprising, therefore, that there are phenomena particularly associated with vehicular stimulation. One example that most people have experienced is the false allocation of motion that can occur when seated in a stationary train as a neighbouring train pulls slowly away from the station. We initially perceive the carriage in which we are seated to be moving, and this false allocation is only later replaced by the veridical percept when the other train is moving faster. Why does this happen? In order to understand it we need to examine the frames of reference that are present. First consider the situation when the train in which we are seated moves out of a station. There is a relative displacement of the platform with respect to the windows of the carriage. Some aspects that were initially visible disappear from view and others that were previously occluded become visible. The rate of these changes increases as the velocity of the train increases. Any stable environment will produce such an optic flow with respect to a moving vehicle, and so it is this characteristic that usually defines vehicular motion. In the case where the neighbouring train starts to move the optic flow produced is ambiguous – it could be a consequence of either train moving, but it is more common for the motion of the whole visual field to specify self motion rather than motion of the environment. The illusion of motion only occurs if the neighbouring train is close enough to occupy most of the view visible through a large window; if it is several tracks away, so that other parts of the station are visible, then there is no illusion. One of the reasons that we cease to experience motion of our own carriage after some seconds is that the illusory motion is not accompanied by the slight carriage motions that do occur when our train is actually moving. We can register these via receptors in the inner ear (in the vestibular system) that signal body motion through space. There is, then, a conflict between visual information signalling self motion and vestibular information signalling that the body is stationary; initially vision is dominant but it is later modified by the vestibular signals for stability. However, the situation is rather more complicated than this, because the eyes rarely remain fixed with respect to the surrounding environment. During the illusory motion we can look around the carriage and perceive that it is stable. Thus, some image motions over the retina (due to eye movements) result in perceptual stability of the carriage whereas others (due to motions relative to the window frame) produce an illusion of motion.

The false motion that is attributed to the carriage in which we are seated is, again, one case of many that can be studied in a laboratory. The experience of self movement as a consequence of visual movement is called vection, and it can occur in the three dimensions of space. That is, we can experience forward body motion when the optic flow is artificially modified to mimic this, or forward tilt or sideways rotation if the visual environment is appropriately manipulated.

These effects are not confined to vehicles and laboratories, as they have a venerable history in the fairground. Many fairground attractions involve abnormal

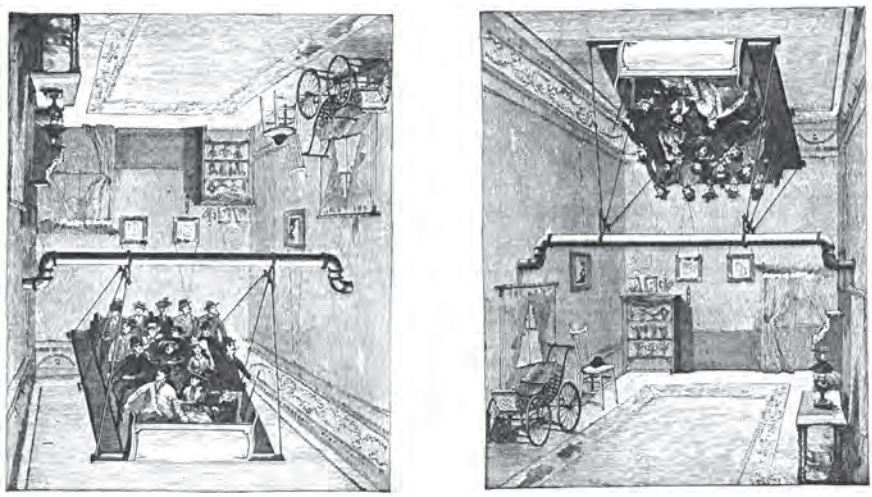


Figure 7.5. An illustration of a haunted swing from the late nineteenth century. The caption for the left picture is "True position of the swing". That for the right is "Illusion produced by a ride in the swing". The two illustrations are the same apart from their orientation.

patterns of motion, both vestibular and visual, and so it is not surprising to find that the desire for abnormal stimulation should have exploited vection, too. There is a venerable and fiendish device called the witch's or haunted swing, on which an individual stands and experiences considerable self motion although little occurs (Figure 7.5). The visible room surrounding the person is not anchored to the ground, but swings about. This induces the impression of self motion in the opposite directions to those of the room; such impressions are intensified when allied with an unstable platform on which the hapless individual stands. The history of the haunted swing, like those of other fairground illusions, is itself mysterious. Since there was advantage in keeping the trick secret, very few records of their construction were made. Thus, in contrast to science in which information is shared, the fairground illusions were guarded secrets. This aspect of secrecy was eloquently described by David Brewster, who we encountered in the context of the lenticular stereoscope. In his *Letters on natural magic addressed to Sir Walter Scott* he wrote: "The secret use which was thus made of scientific discoveries and of remarkable inventions, has no doubt prevented many of them from reaching the present times; but though we are ill informed respecting the progress of the ancients in various departments of the physical sciences, yet we have sufficient evidence that almost every branch of knowledge had contributed its wonders to the magician's budget, and we may even obtain some insight into the scientific acquirements of former ages, by a diligent study of their fables and their miracles".

Sources of motion stimulation

One of the paradoxes of visual motion perception is that movement over the retina is not a necessary condition for its occurrence. This has been established most clearly in laboratory experiments, but it has also been demonstrated in a number of naturally occurring phenomena, as described above. Motion over the retina can be produced in a wide variety of ways, because of the complex motions that observers can make with respect to a stationary environment, as well as the motions of objects in that environment. In this section we will outline the major sources of motion stimulation that follow from observer and object motions.

Observer motions

Observer motions can be considered in terms of the principal articulations that change the position of the eyes in space. The most basic level is that of eye movements. All eye movements relative to the orbit are rotations. The centre of rotation is located about 13.5 mm behind the cornea, and rotations can occur around a vertical axis (towards and away from the nose), around the horizontal axis between the eyes (elevation and depression) and around the optical axis of the eye (torsion). The rotations can be slow and smooth, as when pursuing a moving object, or they can be at very brief and ballistic.

Many eye rotations occur as a consequence of head movements, and their function is to maintain fixation on environmental stimuli despite the change in head position. Such compensatory eye movements result from stimulation of receptors in the vestibular system, and they take place both quickly and with precision; they are referred to as vestibulo-ocular reflexes. Amongst these are the 'doll's eye reflexes' – when the head rotates upward and backward about a horizontal axis the eyes retain the same direction with respect to the environment. That is, they rotate in the opposite direction to the head, as occurs with many dolls which have movable eyes. There are similar compensatory eye rotations for the other rotations of the head in space.

The head undergoes translations as well as rotations in space. The translations can be very complex because of the many articulations that are possible with other body parts – the head on the trunk, rotations of the torso, and bending at the hips and knees. Many of these articulations are a function of locomotion, and the motion of the head during walking and running is more uniform than that of other body parts. This can be seen most clearly in animals with lateral eyes: if you watch the head of a bird, like a seagull, when it is walking, its head remains relatively fixed in space and then moves rapidly forward, remaining there until another rapid forward movement is made. When we walk the head motion follows a path that is more nearly horizontal, despite the appreciable vertical movements of other parts of our body. We do not have the jerky head movements that are common in birds because both our eyes are pointing in the direction of motion. We can fixate on some distant object straight ahead and maintain fixation during locomotion.

The optical consequences of head movements can be distinguished from those produced by eye movements. When a single eye rotates all objects in a scene that are aligned in one eye position remain aligned following a rotation. You can easily ascertain that this is so. Keeping your head still, and closing one eye, look at some objects that are in the same direction but at different distances from you; when you rotate the eye they will remain aligned. Another characteristic of eye rotation is that some parts of the scene that were formerly visible disappear, and others that were not initially visible are disclosed. This occurs symmetrically for elevations and depressions, but not quite so symmetrically for rotations about a vertical axis. The visual field is not evenly bounded on the nasal and temporal sides (see Figure 4.7), and so turning the eye in a nasal direction results in the nose obscuring a larger part of the visual field than is the case when the eye rotates temporally. All these features are unique to rotations of an eye, and they can be contrasted to the consequences of head rotations and translations.

Head rotations, as in turning to the left or right, involve translations of the eye as well as changes in direction. This results in the relative displacements of objects at different distances. Again, you can easily demonstrate this for yourself. Try to keep the position of the single open eye constant in the orbit and rotate the head: objects no longer retain the same alignments, and near objects can occlude parts of more distant ones that were formerly visible or, conversely, disclose parts that were formerly hidden from view. Moreover, the occlusion and disclosure at the left and right boundaries of the visual field does not change the proportion of the field occupied by the nose. Thus, the pattern of visual stimulation can serve to differentiate between rotations of the head and of the eye.

Head movements can also be translations in the three dimensions of space. Forward movement results in a radial expansion of the visual field; if the gaze is straight ahead, then objects at the extremities of the visual field are lost from sight, and distant ones are enlarged. Movement towards a large object results in it occluding more of the background, and the optical expansion increases greatly with impending contact. Sideways movement produces a horizontal optic flow, and upward movement results in a vertical flow. In each case there is a symmetrical change at the extremities of the field, and objects at different distances undergo relative displacements.

Object motions

All the retinal motions described above have been considered with respect to a stable and stationary environment. Additional retinal displacements follow from object motion. When a rigid object moves with respect to a stationary background certain transformations in the optic array occur which can be used to determine the path of object motion. Suppose a rectangular object moves in a frontal plane, then parts of the background will become occluded by the leading edge, other parts will be disclosed by the trailing edge, and there will be shearing along the edges parallel to the motion. You can produce such motion with this book, by holding it at arms length and moving it horizontally and vertically; note the

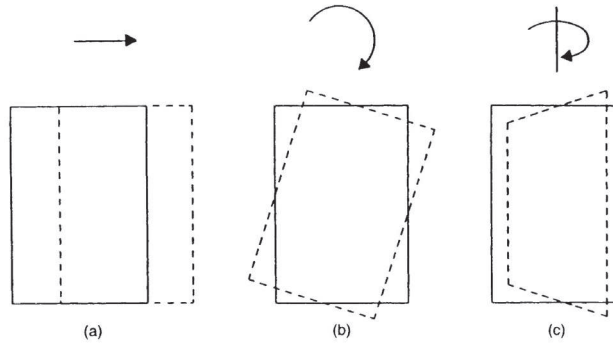


Figure 7.6 Optical transformations produced by a rectangular object (a) moving horizontally in the frontal plane, (b) rotating in the frontal plane, and (c) rotating about a vertical axis. In addition to the optical transformations of the object itself, different aspects of the background are occluded and disclosed by each one.

displacements of the book with respect to objects in the background. This pattern of relative displacements will apply whatever orientation the book is held in, as long as its motion path is in the frontal plane. If, however, the book is rotated in the frontal plane a different pattern of transformations will ensue. Some small part of the background remains occluded continuously, whereas other parts (those in close alignment to the edges of the book) are systematically and symmetrically occluded and disclosed. Yet other transformations follow from varying the angle of the book to the line of sight (rotating it about a vertical axis). The maximum amount of background is occluded when the book is in the frontal plane and the minimum amount when it is in the median plane (with its spine either towards or away from you); between these two extremes, the approaching edge expands optically and the receding edge contracts. Thus, there is optical information available in the patterns of transformation for the paths along which rigid objects move. These three situations are illustrated schematically in Figure 7.6.

In these examples of object translations and rotations we have considered the transformations with respect to the visible background. What would be seen if the background was not visible? This situation is difficult to examine if an object like a book is visible, but it can be studied by attaching points of light to the four corners of the cover of a book and observing them in an otherwise dark room. Motion of the four points in a horizontal direction would not produce any changes in their relative position nor in their relative orientation (see Figure 7.7a). Under these circumstances the motion might not even be detected at all. If it was, it would be due to information from the eyes derived from pursuing the points. This is a situation that we refer to as uniform motion, and it will be discussed in more detail in the following sections. Rotation of the points in the frontal plane would not change the relative orientation of the points – they would remain in a rectangular configuration – but their orientation with respect to gravity would be changed (see Figure 7.7b). We can perceive motion accurately under these conditions, which indicates how important object orientation is in determining our perception

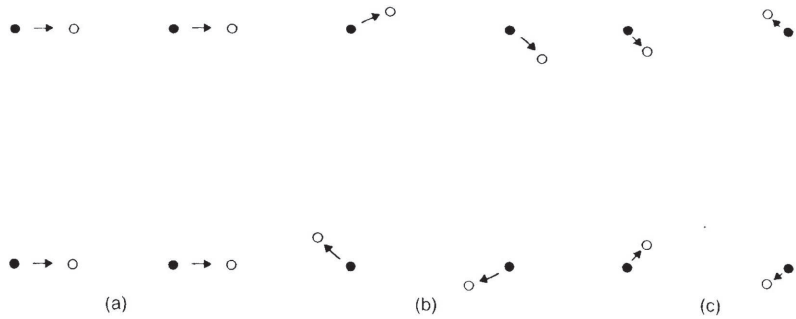


Figure 7.7 Optical transformations produced by four lights at the extremities of a rectangular object (a) moving horizontally in the frontal plane, (b) rotating in the frontal plane, and (c) rotating about a vertical axis

of location and motion. The third transformation, rotation about a vertical axis, modifies both the relative separations and the relative orientations of the points. The lights on the receding edge would remain vertical but their separation would decrease while those on the approaching edge would increase and remain vertical; the separations of the upper and lower points would decrease as would their relative orientations (see Figure 7.7c). These complex transformations are typically seen as a surface changing its bearing with respect to the observer.

These conditions can be examined experimentally but with more naturalistic arrays of light stimuli. Lights can be placed on the joints of moving humans so that the many articulating parts would change the relative positions of the lights. Figure 7.8 illustrates a static view of lights placed on the shoulders, elbows, wrists, hips, knees and ankles of a walking person; not all twelve light sources are necessarily seen because some are occluded by the body. Without some prior knowledge (such as you now have) it would be difficult to discern the nature of the object so illuminated. As soon as the array undergoes movement the object is recognised as a person, and the nature of their movement can also be discriminated – whether running or walking, dancing or drinking.

Our ability to make fine discriminations of biological motions has received considerable experimental attention. With film sequences, human motion can be recognised in as little as 1/10 s, that is, from just two frames of film. Reliable estimates of the gender of the lighted subject can be made, largely on the basis of differences in the centres of moment (the point in the body relative to which all movements are related); it is higher in females than males. With lights on the hips, knees, ankles and toes it is possible to distinguish between a walking male or female from a single step. Observers can estimate the weight of unseen objects on the basis of the relational dynamics of the points on the lifter. Thus, not only is the optic flow to a moving observer transformed in systematic ways, but the articulations of observed humans can be readily perceived. These examples demonstrate the perception of biological motion based on the detection of

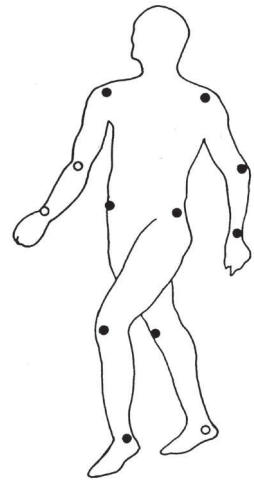


Figure 7.8 Six pairs of lights attached to the major articulations of the body; the filled dots would be visible and the unfilled ones would not be seen in the posture shown here. When a single static array is seen it is not recognised, but as soon as movement is added the lights are recognised as being attached to a person.

particular transformations in the optic array. In general, there is now a large body of work which shows that the perception of many sorts of environmental events and the control of appropriate behaviour can be related to dynamic patterns of visual stimulation, in both man and animals.

Eye movements

When discussing motion phenomena earlier in the chapter, a constant refrain was the attempt to account for them in terms of movements of the eyes. This was so despite the fact that little was known about how the eyes moved until the late nineteenth century. So we can ask, how do the eyes move and what influences the movements? There are many types of eye movements and one of the historical distinctions has been between those that occur within the eye itself and those that move the whole eyeball. The six extraocular muscles were illustrated in Figure 5.18. The muscles turn the eyeball around a centre of rotation. The axes of rotation can be considered as reflecting vertical (for the superior and inferior recti) and horizontal (medial and lateral recti) eye movements and those around the optic axes or torsional movements (superior and inferior obliques). Eye movements serve binocular single vision and they enable extraction of spatial detail. Cones are not distributed evenly throughout the retina; rather their concentration is maximal in the fovea at the centre of the macula lutea or yellow spot, and they decrease rapidly with eccentricity into the periphery (Figure 4.21). This distribution gives rise to distinct vision whereby only the centre of the visual field is resolved clearly and visual acuity decreases toward the periphery. Restriction of high acuity vision

to the centre of the visual field necessitates movement of the eyes in order to bring this window of clear vision to bear on objects of interest. However, the cones do not react to light instantaneously, requiring in the region of 20 ms to respond; consequently, vision becomes blurred when the eyes move. This blurring can be quite severe even when the eye is moving at very slow velocities. These restrictions – of distinct vision and cone physiology – have given rise to an exquisite range of types of eye movement, which operate to satisfy both the need for stability of the retinal image (to allow the cones time to respond) and the need to move the foveas of both eyes to regions or objects of interest.

When we fixate an object in the world, maintaining the image of that object on the visual axis is problematic due to two factors: object and observer movements. A characteristic pattern of eye movement known as optokinetic nystagmus (OKN) occurs when the whole visual scene moves. This happens when looking out of a window in a moving vehicle, and OKN was once known as railway nystagmus for this reason. OKN has two components, called the fast and slow phases. In the slow phase there is smooth pursuit of the moving field, which keeps the image on the same region of the retina. If the velocity of field movement increases above around 30 deg/s, the eyes lag progressively behind, and the maintenance of steady fixation on the object is less effective. The slow pursuit phase alternates with fast, saccadic eye movements that return the eyes to the straight-ahead position. OKN seems to be a primitive form of eye movement control, designed to prevent displacements of the retinal image during head movements.

Experimental interest in eye movements initially arose in the context of vertigo and the influence of the semicircular canals on nystagmus and vestibulo-ocular reflexes (VOR). At the end of the nineteenth century attention shifted to eye movements during reading, and it was in this context that the term ‘saccades’ was introduced. They refer to jerks of the eyes, and it was found, to the surprise of most researchers, that saccadic eye movements were the rule rather than the exception. When reading text, the eyes move with a series of jerks along the line, punctuated by short fixations. These rapid movements were difficult to record and some drastic methods were initially adopted. The system shown in Figure 7.9 is from the 1890s. A plaster of Paris eye cup was placed on the eye (after it had been anaesthetised with cocaine!) and when the eye moved a trace was made on a smoked drum, as shown on the left. The trace displays the saccades and fixations characteristic of reading. More refined photographic techniques were soon introduced which did not involve attachments to the eye.

OKN and VOR operate to maintain a relatively stable image on the retina during fixation, but as we have already stated, the visual system must be able to move the eyes around the scene so that the objects of interest can be projected onto the fovea. In addition, a novel peripheral target or a moving one can elicit a fixation reflex, so that the target falls on the foveas. During saccades, rotational velocities of up to 600 deg/s can be reached. Saccades are a family of ballistic eye movements with durations and peak velocities varying according to the amplitude of the rotation. For any particular amplitude of rotation, saccades show stereotyped position profiles, durations and velocity profiles. The high velocities

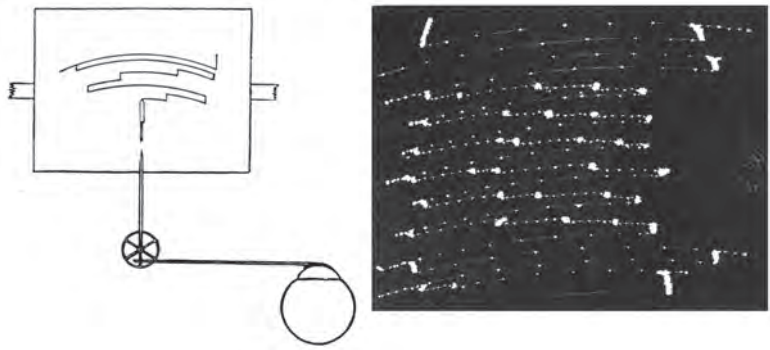


Figure 7.9 Left, one of the earliest technique for recording eye movement during reading. A lever was attached to a plaster eye-cup and moved a pen over a rotating smoked drum. The initial records shown are for calibration across the lines, followed by jerks and pauses during reading. Right, a record of eye movements during reading using the lever device.

reached during saccades ensure that their duration is kept very short ranging from approximately 30 ms for a 5 deg movement to 100 ms for a 40 deg saccade. In this way the period of visual disruption resultant from saccadic eye movements is minimised. During saccades the observer is effectively blind owing to limitations of photoreceptor response time course and to active suppression of the visual pathways during these eye movements. Thus, when we scan a visual scene, we do so primarily using this system of saccadic eye movements, punctuated by periods of fixation that typically last for around 300 ms.

While OKN and VOR work to maintain a relatively stable retinal image during fixation, stability is not absolute; small movements of the eyes do occur during fixation. They undergo high frequency, low amplitude tremors onto which are superimposed slow drifts and rapid flicks or microsaccades. They can be observed by the simple expedient of generating an afterimage and then trying to maintain fixation. By fixating on the central white dot in Figure 7.10 for long enough to generate an afterimage (about 30 s) and then attempting to fixate on the black dot the minor involuntary movements of the eyes will be expressed in shifts of the negative (reversed contrast) afterimage. Such involuntary movements of the eyes are essential for maintaining the visibility of the targets that are fixated. When these minor movements of light from a target over the retina are compensated for then vision ceases after a few seconds.

As primates, we possess one further class of eye movements absent from many other animals; we can pursue a moving target so that light from that target remains on the fovea. If the eye rotates at the same velocity and in the same direction as a moving target then that target will be projected stably to the fovea and so remain clearly visible. While a pursuit eye movement is taking place, the image of the rest of the visual scene is displaced over the retina in the opposite direction to the eye movement. Pursuit eye movements vary in velocity according to that of the target, although there is an upper limit to target velocity of around 30 deg/s if the eyes are to keep up with it.

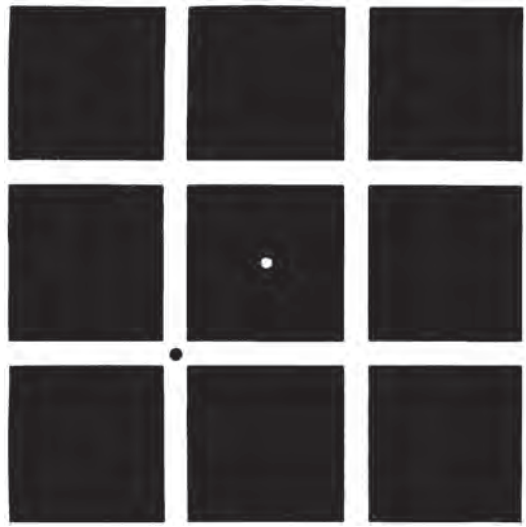


Figure 7.10 A pattern suitable for producing afterimages. Fixate the central white dot for about 30 seconds, under good illumination, then look at a the black dot. You will be able to see the effects of small involuntary eye movements, in the form of rapid displacements of the afterimage squares.

Although the discussion so far has mentioned 'eyes' nothing has been said about their mutual movements. In smooth pursuit of an object at a constant distance from an observer, the two eyes move in the same direction and with the same velocity; these are version movements of the eyes. If the object approaches directly towards the observer then the eyes rotate in opposite directions – they converge, and retain convergence when the object stops. The object is then bifixated – fixated by both eyes. Vergence and version movements of the eyes occur to keep light from a target falling on the foveas of both eyes.

The patterns of eye movements described so far relate to those occurring in normal healthy observers. There are, however, systematic departures from these patterns, and the study of abnormal eye movements has been a cornerstone of medicine for centuries. The most common abnormality is in the lack of coordination between the two eyes so that bifixation is not possible. This occurs in strabismus or squint in which one eye deviates from the direction of the other (see Chapter 2). Strabismus can have early or late onset, and the visual consequences are greatly dependent upon this. For late onset there is initially double vision, but typically one eye is used in preference to the other to achieve singleness of vision. The misalignment of the eyes is very evident to anyone looking at a person so affected, and is probably one of the reasons why medical corrections for the condition have long been sought. Masks for correcting squint have been in use for many centuries (see Figure 2.31). Studies of abnormal movements of the eyes have a similarly long history.

Compensation for eye movements

Eye movements are an inevitable counterpart of normal visual perception, whether to fixate on a different part of the visual field, to keep a moving target on the fovea, or to maintain visibility. Recovery of egocentric movement therefore requires information about both image movements over the retina and the movements of the eyes. In practice, it is only pursuit movements that create a significant problem. Saccadic movements are so fast that retinal stimulation is smeared out, during which there is little possibility of detecting any information. In addition, there may be a suppression of visual signals during a saccade. Image displacements during pursuit are by contrast necessarily similar in extent and velocity to those which occur with object movement. This process of compensation of the retinocentric signal for eye movements was described by Helmholtz, and has since been examined experimentally. In principle, to derive egocentric motion, the requirement is for compensation of retinocentric image motion for pursuit eye movements.

What sources of information are available to the visual system regarding pursuit eye movements? There has been much discussion of this issue, generally focusing on the role of feedback from the extraocular muscles. Muscle groups, such as those that move the limbs, have sensors that respond to being stretched, and can signal the state of contraction of the muscle. This information provides feedback as to whether a movement is taking place as planned. The eye muscles might work in a similar way, and signal their state of contraction, and thus the movements of the eyes to which they are attached. In fact this does not seem to occur, at least as far as movement perception is concerned. A demonstration of this is that when an eye is moved passively, for example by gentle pushing, stationary objects in the environment appear to move. If there was information available from the muscles themselves, this would not happen. The major source of eye movement information is referred to as the efference copy, which means that the signals sent to the eye muscles are stored and also used to represent how the eyes will in fact move. Evidently this can only work if the eyes actually move as they are commanded to do, and it may be faster than an afferent feedback loop. The role of efference in motion perception can be seen with afterimages; when the eyes are moved voluntarily, the afterimage appears to do so also. In this case the absence of retinal motion together with the efference copy corresponds to an object moving at the same speed and in the same direction as the eyes. However, if the eye is moved passively, the afterimage does not appear to move, since there is no voluntary command, and the efference copy therefore specifies that the eyes are stationary. You can confirm this with an afterimage obtained from Figure 7.10. Eye movements can in principle also be signalled by optical cues, as was pointed out earlier. When there is a pursuit eye movement, the whole retinal image displaces with equal velocity and by the same extent, apart from the stimulus that is being pursued. This pattern of change could not normally occur for any other reason, and could provide reliable information about eye movements. To be used, there would have to be some means of detecting and comparing image motion

across the whole visual field. At present, we have little information as to whether there is such a system, and how it might work.

Frames of reference in motion perception

The physical definition of motion is the change of position over time. As we discussed in the last chapter, any statement about position requires a frame of reference relative to which it is measured. Clearly, what is true about static position is also the case for change in position, and therefore motion can also only be described or perceived with respect to a particular frame of reference.

The first frame of reference we will examine is that of the retina itself. Visual motion is mediated by stimulation of the retina, but motion perception is not synonymous with retinal motion. Initially, we will examine the simplified situation of a single eye that is stationary with respect to the head, which in its turn is stationary with respect to the environment. If the eye is fixed in space then object motion would produce motion over the retina. We will consider two conditions of object motion – with and without any background visible. When an object moves in isolation (without any visible background) the only source of information for its motion derives from its displacement over the retina; this is called uniform motion. When a background is also visible, there is information for its uniform motion and also for its displacement relative to the background; this is referred to as relational motion. The threshold for relational motion is much less than that for uniform motion.

Uniform motion

When a target, like a point of light, moves over the retina it stimulates a sequence of receptors, whose local signs provide information about where the stimulation has occurred. In order for the sequence of stimulation to be registered as motion there would need to be some temporal integration between the neighbouring local signs. Visual persistence provides evidence that such temporal integration does occur. That is, if the neural effects of stimulation at neighbouring local signs outlast the duration of stimulation, then they will be active at the same time even though the physical stimulation is successive. Temporal integration only operates within certain limits: if the motion is too slow it will be registered as successive; if it is too fast it will be registered as simultaneous. The same retinal motion signal would be produced by different motion paths of objects in the external environment, provided that the angular velocity with respect to the eye was the same. For example, two objects might be at different distances, but the nearer one could be moving more slowly than the farther one. If the distances and velocities are chosen appropriately, the retinal images of both would move across the same distance on the retina in the same time. The retinocentric motion signal is the information available to the visual system about the displacement of retinal images with respect to the retina.

The retinocentric frame of reference is not confined to the retina, but is constrained by the retinal coordinate system. Neurophysiologists have demonstrated that the cells in the visual cortex are coded retinocentrically; that is, they respond when particular regions of the retina are stimulated, regardless of the direction the eyes are pointing. Many cells in the visual cortex are specifically responsive to motion in a particular direction and also at a particular velocity. For example, one cell might respond to a horizontally oriented edge moving downwards over a particular region of the retina; it would also respond to a stationary horizontal edge in the same region, but not as strongly as to the downward motion; upward motion of a horizontal edge over the same region would have very little influence on the cell's activity. Another cell can have the opposite pattern of activity. In this way, different cells will respond to all orientations and directions of motion. Cortical cells could only respond in these ways if there is temporal integration over retinal local signs.

Considerable attention has been directed to understanding the ways in which the visual system might implement retinocentric motion detection. A key question concerns what is, or could be, detected. Under normal circumstances, when a real object moves, all its properties move with it. Its shape, colour, distance and luminance displace together, because these are all properties of the object as a single physical entity. Similarly, the object's retinal image, which carries information about these physical properties, moves in a coherent manner. It is, however, a matter of empirical investigation to determine which of these properties is used by the visual system to detect motion. The most fundamental property is luminance. Consider a simple luminous square against a black background. Its retinal image will produce a response from the area of retina on which it falls. If it moves to another location in space, and the eyes are stationary, then there will be a change in the location of retinal stimulation. Another way of expressing this is to say that there has been a change in the spatio-temporal distribution of light energy on the retina. Various systems have been proposed that could pick up this change. The simplest sort of motion detector would link the luminance detecting elements at the areas of retina stimulated during the motion. If the output from the luminance detectors activated by the square at its initial position were to be delayed, and subsequently combined with the signal from the square at its final position, then when both coincide, motion has taken place between the two retinal locations. A cell that received the delayed and concurrent inputs of such luminance detectors would function as a simple motion detector. Of course, the delay would have to correspond to the retinal velocity of the stimulus, so motion detectors with different delays and/or separations would be tuned to detect different velocities. This simple scheme can be improved and elaborated on by various sorts of additional processing, but the principle of operation is unchanged. It is noteworthy that such a motion detector has a side-effect: it will signal motion if there is no continuous movement between two retinal locations, but instead a stimulus is briefly presented at one point and then briefly presented at another. This is the basic requirement for the perception of apparent movement in a series of briefly presented static scenes, as described above.

This type of motion perception is often termed 'first-order' or 'Fourier' motion, because of the presence of spatio-temporal luminance change. As with the perception of depth from retinal disparity, there is evidence from random dot displays that object recognition is not necessary to detect first-order motion. A random dot cinematogram, in which an area of texture is displaced with respect to a textured background, is the logical equivalent of a random dot stereogram. When the textured area is moving, it is seen clearly as a distinct shape. When the movement stops, no shape can be seen. This corresponds to the perception of a shape separated in depth from the background in a random dot stereogram. In this case, one eye sees an area of texture which is displaced with respect to the equivalent area seen by the other eye, but no shape can be seen using either eye alone. There are very interesting parallels between motion and stereo perception, perhaps reflecting the evolutionary derivation of the latter from the former. All visual systems detect motion, but only some animals have stereoscopic vision. Possibly, as is often the case in evolution, a pre-existing system became adapted for a new purpose.

However, there is also good evidence that we can see motion when there is no displacement of a luminance boundary on the retina. This is termed 'second-order' or 'non-Fourier' motion. Suppose that a random dot field consists of dots which are constantly changing from black to white and vice versa; that is, the whole field appears to twinkle. If a portion of the field is displaced, the mean luminance is everywhere the same, and there is no moving luminance boundary. Nevertheless, observers can see motion in the field. The same is true of moving boundaries defined by colour or contrast instead of luminance. Simple motion detectors could be used to drive systems which apply transformations to the first-order signals, and in theory second-order motion perception could be explained in this way. Alternatively, observers may pay attention to features defined by colour or contrast, and this attentive tracking has been the subject of much research. Probably both these processes play a part in such motion perception, but natural circumstances which might involve non-Fourier motion alone are quite difficult to imagine.

Some visual motion phenomena, like the motion aftereffect, have been interpreted in terms of processes like simple first-order motion detectors. Consider what happens when an observer initially looks at the rocks by the side of a waterfall: the stationary rocks will have many contours which will stimulate physiological edge detectors. Many of these edge detectors will respond more strongly to motion in one direction than in the opposite. For example, horizontal contours will excite motion detectors for downward and upward movement, but the net effect of these would cancel. Thus, the perception of stationariness is dependent upon the balanced activity of motion detectors coding opposite directions of motion, as shown in Figure 7.11. When the waterfall is observed the downward motion detectors will be strongly stimulated; if this stimulation is prolonged, the motion detectors will adapt or fatigue. Subsequent observation of the stationary rocks will produce a different net effect: the fatigued downward motion detectors will exert less influence than the unadapted upward motion

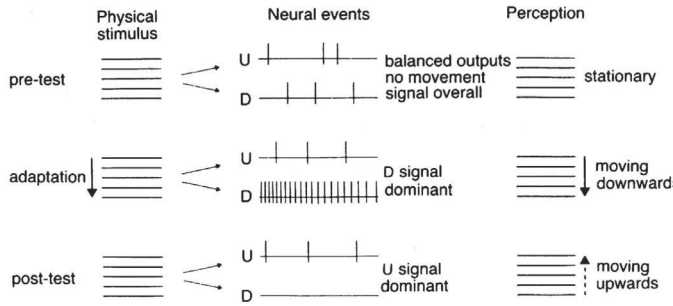


Figure 7.11 An interpretation of the motion aftereffect based upon adaptation of cortical motion detectors. In the pre-test a stationary grating will excite motion detectors for upward (U) and downward (D) directions equally, and so the net activity will favour neither, resulting in its stationary appearance. During adaptation the downward direction detectors are strongly excited resulting in a neural signal for that direction, which is reflected in perception. In the post-test the stationary grating is presented again. The downward direction detectors have been adapted due to prolonged stimulation and so are not responsive. The upward detectors signal as before, but since there is no balancing output from the downward detectors the net activity signals upward motion, and the stationary grating is seen as ascending.

detectors. Therefore, the signal from the stationary rocks would be similar to one produced by contours moving slowly upwards, and that corresponds to what is seen. The effect is considered to be retinocentric because the motion aftereffect is confined to the region of the retina that has been exposed to the motion. However, there is evidence that second-order motion can give rise to a motion aftereffect, and also that the first-order (luminance change) type is influenced by quite complex features of the surrounding configuration. The latter are discussed below. It seems likely that there are several motion aftereffects, not one, and that they reveal different motion detecting processes, together with the influencing factors to which each is subject.

Recordings from single cortical cells fulfil the requirements for retinocentric stimulation because the experimental animals are anaesthetised and the eye is immobilised. Therefore any stimulus motion will have a geometrically corresponding retinal image motion. This rarely occurs with natural vision because of eye movements. However, stroboscopic motion does come close to meeting the requirements for retinocentric stimulation: apparent motion can be seen between two brief, stationary and spatially distinct pulses of light, if the temporal and spatial separations are appropriate. Two different types of stroboscopic or apparent motion have been described. One has been called the short-range process because it is confined to spatial separations between the stimuli of less than 15 min of visual angle, and temporal intervals of about 100 ms. The short-range process can only be activated if both stimuli are presented to the same eye. The long-range process is elicited by wider separations and longer intervals, as well

as occurring with stimuli presented to each eye in turn. This distinction has been questioned by later workers and it may not reflect genuinely different underlying processes. Nevertheless, it has been very influential in the development of ideas about motion detectors, and in determining the types of experiments on apparent movement carried out over many years.

Relational motion

Returning to the motion aftereffect, waterfalls are not seen in isolation; they occur in a wide variety of terrains. Accordingly, waterfall illusions in the natural environment involve relational motion. The water descends with respect to the "age-worn rocks immediately contiguous", and so there is an alternative visible frame of reference – a patterncentric one. Motion adaptation examined in the confines of the laboratory is similarly dependent upon relational motion. Rotating sectored discs and spirals present relational motions between their individual parts, but aftereffects based on linear motion require a patterned background relative to which the motion can be allocated. Therefore, even one of the phenomena that was thought to be dependent on uniform motion is in fact using patterncentric signals, as is evident from the experiment illustrated in Figure 6.2. It might be expected that a large uniformly moving display, like a vertically striped cylinder rotating around an observer, would be the ideal stimulus for generating a motion aftereffect. In fact it is not; rather than producing a visual motion aftereffect it results in the observer's apparent self motion. It is an example of vection, similar to the situation in the train mentioned earlier in the chapter.

Patterncentric signals register the relational motions between stimuli, rather than the retinocentric motions. One of the targets is adopted as a frame of reference relative to which the motion of the other target can be registered, and vice versa. Thus, patterncentric frames of reference are confined to interactions within the pattern projected onto the retina. The patterncentric signal remains the same even if the eyes move, because this does not alter the object separations. Eye movement simply adds a constant amount to the displacement of all retinal images, leaving differences in displacement unchanged. There are as many patterncentric frames of reference available as there are independently moving objects. Although this might sound a dauntingly complex process for the visual system to undertake, it is not so in practice because many of the patterncentric frames of reference are equivalent. Consider the eye fixed in space and viewing the waterfall. The patterncentric frame of reference could be taken as the descending water or as the rocks. In the former case the water would be taken as stationary and the rocks would be registered as moving upwards, or the rocks could be taken as stationary and the water descending. We do not need to use knowledge of the world to adopt the latter frame of reference, because there is much more correlated patterncentric information for the stability of the rocks rather than the water. Many other objects projected onto the retina, like trees, grass, or the river bed, yield the same outcome as that for the rocks, namely that the water is descending and the other features remain in the same relation to one another. Under normal circumstances, the

environment provides a stable patterncentric frame of reference relative to which object motions can be allocated.

The night sky is somewhat unusual in this regard, and the induced movement of the moon provides us with a phenomenon that can be used to emphasise the points made above. Induced motion can be studied in the laboratory with displays like those shown in Figure 7.1, which allow us to study the uniform and relational components in retinocentric motion. One of the simplest displays consists of three points of light in an otherwise dark room; the central point remains stationary and the outer two move in the same direction at the same speed. Under these circumstances observers report that the central point appears to move in the opposite direction: motion is induced in it. The two moving points provide a more powerful patterncentric frame of reference than the single stationary one. On the one hand, if the patterncentric frame was the only factor operating then the two outer points would not appear to move at all; on the other, if the uniform process operated alone the central point would appear stationary and the outer points would appear to move. However, what typically happens is that motion is seen in both components of the display. This outcome indicates that both uniform and patterncentric processes are implicated in induced movement. The same outcome applies to the other displays shown in Figure 7.1.

Up to this stage we have considered motion with respect to a single eye. Relational motion can also be considered to operate between the eyes. That is, different patterns of retinocentric stimulation can occur with respect to each eye. This can be illustrated in terms of a second type of stroboscopic motion, the long-range process. It is possible to present the first stimulus to one eye and the second to the other and apparent motion still occurs. This technique of splitting the stimulation between the eyes is widespread in vision research. It is called dichoptic presentation, and it can be used to indicate whether a particular phenomenon is based upon neural processes at or before the stage of binocular combination or at or beyond it. The signals from the two eyes are first combined at the level of the visual cortex (see Figure 5.5). Therefore, if a phenomenon cannot be elicited when the components are presented dichoptically, it is likely that it is due to neural processes occurring before the visual cortex. Alternately, when the phenomenon (like long-range apparent motion) does occur with dichoptic stimulation, it is suggestive of a more central site for its occurrence, at or beyond the level of binocular combination.

Induced motion has also been examined dichoptically. That is, the inducing stimulus can be presented to one eye and the stationary stimulus to the other. The results from such experiments are not easy to interpret because some have shown dichoptic effects and others have not. The situation is difficult to examine experimentally because there is no common stimulus to keep the eyes in alignment. When induced motion has been produced under dichoptic conditions it is possible that the eyes moved with respect to one another, and so the motion seen could have been due to uniform motion alone.

Another popular technique that addresses a similar question is interocular transfer. The motion aftereffect can be elicited if one eye is used for observation

and if both eyes are used. What would be the outcome if one eye observed the moving display and the other eye viewed the stationary test? The aftereffect still occurs, but it does not last as long as when the same eye is used for both inspection and test. If the effect was confined to monocular processes alone it would not show interocular transfer; since it does, this suggests that binocular processes are also implicated in the motion aftereffect. Typically, the magnitude of interocular transfer in aftereffects is around 50 per cent. The value is about the same for the tilt aftereffect as it is for the motion aftereffect, which suggests that it might represent a fairly general feature of the way spatial information is coded in the visual system. The involvement of binocular processes in spatial aftereffects is supported by the finding that individuals lacking stereopsis (usually as a consequence of an uncorrected squint in childhood) also fail to experience any interocular transfer of aftereffects. It has been found that much larger values for interocular transfer of the motion aftereffect (up to 100 per cent) can be obtained when the test field is flickered at 2–3 Hz. The reason for this remains unclear, but the finding is hard to relate to existing theories and may open up new experimental approaches to this long-standing issue.

Egocentric motion

Human eyes are only stationary when they are paralysed or damaged. As a result, by far the most common cause of retinal image motion is the movements of the eyes themselves. Retinocentric motion information is therefore ambiguous, since it could arise from any combination of object movement and eye movement. In order for the visual system to allocate the retinal image displacement between object movement and eye movement correctly, there must be a source of information about eye movement. Only when both image displacement and eye movements are considered together is it possible to recover changes in the egocentric direction of an object in the environment. The outcome is a representation of object movement with respect to the self; that is, egocentric motion, which is independent of the motions of the eyes in the head.

The account of egocentric motion so far has been based on the assumption that image motions and eye movements are represented correctly in the visual system. Clearly, there might be many reasons why this would not be so. Detection of image motion, like any other sensory process, will depend on the stimulus exceeding a threshold value, and this may vary across the retina. Sensitivity to motion is better in the fovea than in the periphery, although motion is more likely to attract attention in the latter case. Similarly, the signal representing eye movements may not correspond to the actual value. When such mismatches occur, then there should be predictable errors of motion perception. Some examples were given above for active and passive eye movements, but these represent extreme cases. In general, if the extent of eye movements is under-represented, a visible object will appear to move by more than it actually does, and the converse applies to image motion. Illusions like this can be observed. If a moving object, like a pencil, is pursued by eye movements across a patterned background, the background appears to

move in the opposite direction. This can be explained if the internal value for eye movement is too small, so that subtracting it from the image motion of the background gives a non-zero result. If so, it would be predicted that people should underestimate the velocity of the pursued target. This underestimation does occur, and is known as the Aubert–Fleischl effect after its first investigators.

Geocentric motion

The discussion of egocentric motion was concerned with displacements of the retinal image due to movements of the eyes in the head. Such eye movements are rotations, and can therefore be directly related to the angular extent of image motion. Compensation for the effects of eye movements can be based on simple representations of the extent of rotation of the eyes, and the extent of image displacement on the retina. However, retinal images may move as the result of head movements, which also change the location of the eyes in space. Eye movements occur with respect to a frame of reference provided by the head, but movements of the head occur in three-dimensional space. The head may turn with respect to the shoulders, but the centre of rotation is not the same as that of the eyes. A more common situation is that the head, and the eyes, move to different spatial locations as the result of movements of the whole body. These may be active, as in locomotion, or passive, as in travelling in a vehicle. These sorts of movement take place with respect to the physical environment, which therefore provides the frame of reference. Movement expressed with respect to the environmental frame of reference is geocentric. A change in egocentric direction as the result of a head movement cannot be interpreted unambiguously unless the extent of both the head movement and of object motion in three dimensions are known. A further stage of analysis, incorporating information for both self movement and perceived object distance, is therefore required. The geocentric representation which results is the basis for motion perception.

Perceived distance and motion

As an example of a phenomenon which demonstrates the interrelation of distance and movement perception, try folding a strip of paper into the shape illustrated in Figure 7.12. Note that the illustration is not the basis for the effect; it is necessary to observe the folded paper itself. It is usually called the Mach card, and once it is prepared, place it on a table top in front of you under even illumination, and look at it with one eye, by covering the other eye with your hand. After a minute or so you should find that a striking perceptual change occurs, and that there is an inversion of the shape of the folded paper. The peaks become troughs, and vice versa, which means that their perceived distances no longer correspond to their physical distances. Once you have seen this configuration it will be easier to obtain in future, so it is worth persevering. You may notice changes in the apparent brightness of the illuminated and shaded surfaces, indicating that the interpretation of surface luminance is also dependent on distance information. More obviously,

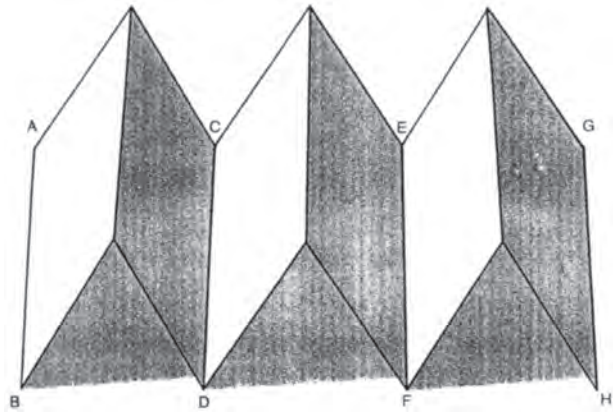


Figure 7.12 Mach's card. How to fold the paper strip for the demonstration is described in the text. The paper should be placed on a flat surface so that all the labelled points are touching it. Observe with one eye, and a stationary head. After a time you should see a change in the apparent orientation of the strip. Points A, C, E and G might appear to be in the air, and B, D, F and H on the surface. AB, CD, EF and GH will appear vertical and nearest to you. Small movements of your head will create striking deformations of the shape of the strip. Mach's folded card demonstrates the interplay between perceived distance and motion, although Mach was concerned principally with variations in perceived brightness rather than motion.

if you move your head slightly from side to side (large movements will destroy the inversion) you will see large shifts in the position of the folds. They will probably seem to move as if the paper is elastic, and being flexed from side to side, with the apparently nearer edges moving in the same direction as your head, and the apparently further ones in the opposite direction.

This perplexing interaction between perceived distance and motion has been exploited by the artist Patrick Hughes. He has produced paintings in reverse perspective; they are painted on protruding planes (like truncated pyramids and wedges) so that the parts that are physically close to the observer are pictorially distant. That is, the lines that would converge on a flat picture plane to allude to distant objects are physically closer in these works. They appear as flat paintings until the observer moves whereupon they undergo a plastic motion that is beguiling. The pictorially distant (but physically closer) parts move in the direction opposite head movement. This outcome that would be predicted on the basis of experiments on concomitant head movement and perceived distance: if the perceived distance is greater than the physical distance then a retinally stationary stimulus will appear to move in the opposite direction to the head. When standing close to the works they can switch between the two states, which could be called pictorial and veridical, and equivalent lateral head movements produce dramatically different apparent motions. In the pictorial state the corners appear fluid and in motion whereas in the veridical state there is little apparent motion. If the eyes have retained the same convergence during these two states then the apparent change cannot be attributed to the pattern of retinal motion

during lateral head movement. The same effects can be seen in Mach's card when it is constructed, but not in a drawing of it as in Figure 7.12. They are more immediate and compelling with a painted model.

Similar interactions between perceived distance and motion can be observed with stereograms, since they also appear to move when the observer does. Points which are seen with crossed disparity, nearer than the image plane, will move with the head, while points with uncrossed disparity, that appear further away, seem to move against the head. These examples show that perceived distance directly influences the perception of motion. The explanation lies in the processes required to recover geocentric motion from egocentric motion. Consider a stationary object located at a constant distance. For a stationary observer, the egocentric motion will be zero, if changes in retinal stimulation due to eye movements are fully compensated for. But if the observer's head moves, there will be a shift in egocentric direction, whose extent depends on the distance of the object. Thus, a given head movement produces different retinal motions, and therefore egocentric motions, depending on how far away the observed object is. If the object is to be correctly seen as geocentrically stationary, despite head movements, then a representation is needed of its position in three-dimensional space. This can be obtained from the egocentric motion by scaling it according to perceived distance. If the perceived distance matches the physical distance, the scaled change in egocentric direction will match the extent of the head movement. However, if the perceived distance is in error, geocentric motion perception will also be in error. Figure 7.13 shows the effects of underestimation and overestimation of perceived distance on the apparent motion of a physically stationary object. The object may appear to move either left or right, and by varying amounts, depending on the error in perceived distance. Clearly, stimulation on the retina is wholly inadequate to predict what will be seen.

The same analysis can be applied if an object is also moving during observer movement. As with a stationary object, its geocentric motion can only be obtained if its egocentric angular motion is scaled by its perceived distance. Errors in perceived distance could cause an object which is moving to the left being seen to move to the right, and vice versa. The various cues to perceived distance have been discussed in Chapter 6, and there are many situations where errors can occur. A good example is provided by looking out of a window that has marks or dirt on the glass. These are nearby, and there are good cues to their distance. Because of the equidistance tendency, objects seen through the window will appear to be in the same plane as the marks, and therefore appear to be nearer than they are. Consequently, if the head is moved laterally, objects beyond the window appear to move in the same direction as the head.

Perceived self motion

Our discussion of perceived distance in motion perception makes clear that information is also required about self movement. The egocentric motion signal with respect to a subordinate frame of reference (the head) needs to be

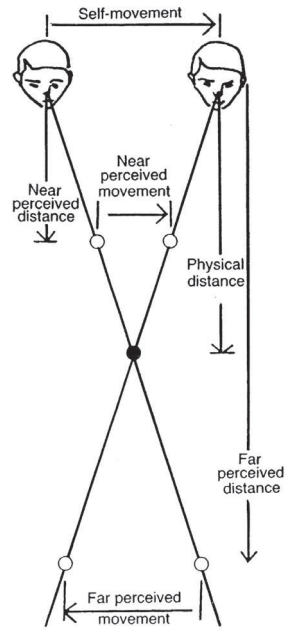


Figure 7.13 The relationship between the perceived position of an object, perceived distance, and self movement. An observer moves his head through a distance towards the right, while observing a stationary object, represented by the filled circle. If the object is seen at its physical distance, it will not appear to move. However, if there is an error of perceived distance (open circles), then the object will be perceived as moving. If it is seen as nearer than it really is, it will appear to move in the same direction as the head; if it is seen as further away, the perceived movement will be opposite to the head movement. The extent and direction of apparent movement during self movement can therefore be used as an indirect measure of perceived distance. The same relationships apply if the object is physically moving; it will only be seen correctly during self movement if it is seen at the correct distance.

compensated for the effects of head movements if geocentric object motion is to be obtained. Since head movements are three-dimensional, egocentric object movement must be expressed as a displacement in three-dimensional space before the two can be combined. An internal value for the extent of self movement plays the same role at the geocentric level as does the value for eye movement at the egocentric level. As might be anticipated, if perceived self motion differs from that which actually takes place, errors in perceived object motion will result. There are a number of sources of information which contribute to the sense of self motion. The vestibular system in the inner ear detects acceleration in each of the three spatial axes. The three semi-circular canals are fluid-filled tubes oriented at right angles to one another, so that a change in velocity of the head results in a movement of the fluid within the canals most nearly aligned with the direction of motion. Fluid movement is detected by the displacement of hairs embedded

in cells in the canals. This system detects only angular accelerations, and does not respond to a constant velocity of movement. Associated structures called the otolith organs detect the static orientation of the head with respect to gravity, as well as other linear accelerations. Sensory systems in the muscles, tendons and joints provide information about posture and limb movement. It is also likely that there is an efference copy of motor commands to the limbs which functions similarly to the efference copy of eye movements.

The function of the vestibular system, and its role in motion perception, is shown by the effects of prolonged stimulation. If you spin round on the spot, or are spun round on a rotating chair, there is strong stimulation of the vestibular system from the rotary motion. As is well known, there are marked aftereffects of rotation. These may include nausea, and you should only try to experience the effects yourself with care. The fluid in the semicircular canals continues to circulate for a time after rotation has ceased, giving rise to a stimulus which corresponds to continuing rotation in the opposite direction. This causes observers to feel as if they are continuing to rotate, even though they are stationary. This illusory self motion is associated with an illusion of object motion called the oculo-gyral illusion. Stationary objects seen in isolation appear to be moving in the same direction and at the same speed as the observer's body. In addition, there is a characteristic pattern of eye movements, referred to as postrotary nystagmus. This is similar to OKN, and consists of alternating fast and slow phases of eye movement, as if the visual scene were actually continuing to move. It is an interesting historical fact that the characteristics of postrotary nystagmus were clearly demonstrated before the function of the vestibular system was understood.

Vision itself also contributes to the perception of self motion. This process, which Gibson termed visual proprioception, depends on the detection of characteristic transformations in the optic flow, such as the radial expansion associated with approaching a stationary surface. If the whole visual field, or a large part of it, changes in a defined manner, then this can be used to derive changes in posture. In effect, the motion of the visual field with respect to the observer is interpreted as motion of the observer with respect to a stationary environment. This is almost certainly correct if the transformation affects most of the visual field. We, and many animals, seem to make this assumption automatically, as is shown by the witch's swing example mentioned earlier. The advantage of visual proprioception is that it can signal constant velocities within the range of biological motion, which the vestibular system is unable to do. In laboratory studies of vection, the sense of self movement takes some time to build up. The initial rotation of the surrounding drum involves acceleration from rest, and the visual signal conflicts with the vestibular signal. The latter takes precedence, and the drum is seen as moving. However, when the drum has reached a constant velocity, there is no sensory conflict. The signal from the vestibular system will be that there is no acceleration of the self, and this is compatible with moving at a constant velocity. Hence the observer experiences moving at a velocity determined by the visual motion. This is an illusion, in that the observer is actually stationary. As a result, if an image which is stationary with respect to the environment is superimposed

on the rotating drum, the stationary image is seen to move in the opposite direction to the drum. The geocentric motion of the stationary image is obtained by compensating the egocentric motion signal (zero) for perceived self motion (positive but incorrect).

When we travel in moving vehicles, the landscape outside often appears to be in motion. The cause of this is probably the lack of sufficient kinaesthetic information for self movement. Most vehicles have some form of cushioning or suspension, so vestibular stimulation is limited except when there is marked acceleration or deceleration. The optic flow may contribute some sense of motion, but perhaps because the velocity is higher than could occur from self-produced motion, this is not sufficient to override the vestibular signal for being stationary. Consequently, the egocentric motion of the landscape is mostly perceived as geocentric motion, and the world appears to be moving. The fact that we know that this is not the case may allow us to act appropriately, but does not change the perceptual experience.

Our discussion of motion perception has reviewed some of the many visual phenomena associated with motion, and a model for understanding how we perceive veridical motion despite the movements of our eyes, heads and selves was proposed. While there have been clear advances in understanding the early stages in motion perception, there is still much to be added. There is a great deal of evidence that the displacement of an image over the retina, whatever its qualitative or quantitative spatio-temporal energy distribution, is not enough to specify what motion we will actually perceive. Retinal image displacement can be reduced, nulled or even reversed in perception by manipulating the perceived distance of the moving object. The pattern of retinal stimulation could therefore vary greatly and still produce the same perceptual outcome. Our present neurophysiological models of retinal motion detection will need considerable elaboration, or even replacement, if a full understanding of motion perception is to be achieved.

Reference notes

All vision is motion vision according to Gibson (1968), and so it is no surprise that motion perception is a topic covered by all textbooks on vision and that some books are dedicated specifically to it (see Epstein and Rogers, 1995). The discovery of brain regions specialising in motion processing has added to the intrigue of interpreting how we see movement and its converse – how we see a stable world despite our eye and body movements. Mather's (2008) book provides an excellent introduction to motion perception, as do Purves and Lotto (2010). Frisby and Stone (2010) describe the physiological underpinnings of motion processing as well as computational approaches to motion detection. The basic phenomena of visual motion described in this chapter have long histories, accounts of which can be found in Boring (1942) and Wade (1998a).

Empiricist philosophers have been concerned with the perception of motion because it did not correspond to any simple pattern of stimulation; Locke's quotation regarding motion and space perception is from his *Essay* (1690). Induced motion

was recorded by Lucretius in his compendium *On the nature of things*; many translations of this work are available and the quotation is from Lucretius (1975). The apparent motion of the moon when clouds move past it was clearly described by Ibn al-Haytham in the tenth century (see Sabra, 1989), and the quotation given can be found in Porterfield (1759). The considerations of induced motion by Lucretius were also linked to the motion aftereffect; both ancient and modern aspects of which are presented in Mather, Verstraten and Anstis (1998). Plateau's description of the spiral aftereffect was given in 1849. The somewhat perjorative quotation regarding the aftereffect from Plateau's spiral and eye movement theories was from Holland (1965), although he used the spiral aftereffect mainly as a clinical test. Despite its antiquity, the motion aftereffect continues to fascinate students of vision, and it has the facility of bending with the fashions of visual science. The same cannot be said of autokinetic sensations, since the perceptual phenomenon was initially confused with a physical one. Humboldt's initial description was in 1799 but his published account did not appear until later; the quotation is translated from Humboldt (1850). Aubert (1887) added enormously to the study of visual motion generally as well as naming the autokinetic sensations. Stroboscopic motion has a longer and more diverse history, since it is related to a phenomenon of antiquity (visual persistence) and it is expressed in a modern enterprise (television and cinematic motion). Accordingly, it is rich in books on its history, of which Mannoni (2000) and Herbert (2000) are amongst the best. Faraday's quotation that stimulated others to make revolving discs is from his article on a peculiar optical deception (Faraday, 1831). The instruments enlisted to supply stroboscopic motion, taken from the wonderful collection assembled by Werner Nekes, are illustrated in Mannoni, Nekes and Warner (2004). The illustration of the phenakistoscope was after one in Helmholtz (1925/2000). In the nineteenth century, instruments of this type were called philosophical toys because they served the dual function of assisting natural philosophers with their experiments and providing a source of popular amusement. A history of philosophical toys can be read in Turner (1998) and Wade (2004). The study of kinetic depth effects is more recent, and they represent a phenomenon that has proved attractive to artists. Wallach and O'Connell gave them this name in 1953, although Musatti (1924) had examined them before that. Musatti call it the stereokinetic phenomenon and Zanforlin (2006) has provided an account of research on it. Marcel Duchamp's rotoreliefs followed rapidly on the heels of Musatti's experiments with the stereokinetic effect. Duchamp produced twelve rotary designs between 1925 and 1935 (Guigon and Pierre, 2005). He maintained an active interest in optics and motion, and the rotoreliefs provided him with evidence of the ways in which vision can be confounded. Examples of such rotary effects can be found at: http://www.michaelbach.de/ot/mot_ske/index.html and http://lite.bu.edu/vision-flash10/applets/Depth/Benussi/Rotorelief.html as well as http://www.opprints.co.uk/stereokinetic-3-4-6.php .

Wells' (1792) experiments on vertigo were published in a book he wrote on binocular single vision. The quotation is from this book which has been reprinted in Wade (2003), together with Wells' other insightful articles on visual vertigo.

Wells' work has been neglected and his experiments were unknown to Purkinje who essentially repeated them decades later. Purkinje did influence the work of many others, including Mach, whose book on the movement sense has been translated into English (Young, Henn and Scherberger, 2001). The experiments by Wells can also be considered as providing the foundation for experimental studies of eye movements (see Wade, 2010). Rotating the body results in the world appearing to spin around the observer. Similar effects can be produced by rotating the visible world relative to a static observer and this produces the sensations of body rotation. This is generally achieved by rotating a striped cylinder round a seated observer. Howard (1982; Howard and Rogers, 2012) provides a thorough account of vection. The illustration of the haunted swing is from Hopkins (1898); it must have been like the swing employed at the Midwinter Fair held at San Francisco in 1895 which could hold more than 40 people!

Before the onset of cinematography, sequences of images from moving observers with lighted limbs were investigated by Étienne-Jules Marey (see Braun, 1992; Mannoni, 2000). It was called chronophotography and was the forerunner of studies of perceived biological motion with point light sources. The modern studies began with the work of the Swedish psychologist Gunnar Johansson (1973, 1975) and have been extended to perception of activities and characteristics other than walking (Mather and Murdoch, 1994; Thirkettle, Benton and Scott-Samuel, 2009).

Understanding of the nature and control of eye movements has advanced very rapidly in recent years, as has the technology for their measurement. Wade and Tatler's (2005) book surveys the historical development of methods for the measurement of eye movements, and sets the scene for the current explosion of research. *The Oxford handbook of eye movements* (Liversedge, Gilchrist and Everling, 2011) provides a comprehensive overview of the wide-ranging scope of current work and its antecedents. The measurement of eye movements with a freely moving observer has taken investigation outside the laboratory (see Land and Tatler, 2009) and begun to provide understanding of the interplay between eye, head and body movements in an active perceiver. Perceptual stability despite eye movements is reviewed by Bridgeman (2010).

The analysis of frames of reference in motion perception presented here corresponds to that used in previous chapters for the perception of orientation and direction. The key concept of hierarchical frames of reference and the specification of the information required to convert one to another (e.g. retinocentric to egocentric by incorporating eye movement information) was developed over a number of years, initially for uniform motion, but later extended to patterncentric motion (Swanston, Wade and Day, 1987; Swanston, Wade and Ono, 1990; Wade and Swanston, 1987, 1996). Subsequently, the ideas developed in these publications have been employed in our experimental work on motion perception, especially the motion aftereffect (Swanston, 1994; Wade, Spillman and Swanston, 1996; Wade, Swanston and de Weert, 1993). A Bayesian model of the Aubert–Fleischl phenomenon has been advanced by Freeman, Champion and Warren (2010).

The relationship between perceived size and movement with perceived distance and self motion has been most fully analysed, theoretically and experimentally, by Gogel (e.g. Gogel (1990) and see also Chapter 8 Reference Notes). Patrick Hughes refers to his works as reverspectives and examples of them can be found in Slyce (1998). Papathomas (2007) has examined reverspectives from a perceptual point of view, and an animated version of a Hughes work can be seen in Wade and Hughes (1999). Mach's (1866) folded card demonstration was concerned principally with variations in perceived brightness rather than motion. When the depth of the card does reverse, any lateral head movements result in marked motions of the peaks and troughs. However, the Mach card continues to oscillate in depth whereas the painted works of Hughes appear to remain stable and pictorially flat when viewed from other than very close proximity. Mach's account has been translated by Ratliff (1965).

8 Spatial representation

Throughout this book we have emphasised the applied, functional character of vision. Vision exists in humans and other animals because it is an aid to survival. The properties of vision are those needed to carry out significant activities such as seeking food, shelter and sex, so that individuals can survive and reproduce. That is, vision and the other senses have evolved to extract information regarding the constant properties of objects despite the changing patterns of stimulation at the sensory surfaces. Thus, object constancies are at the very heart of perception and in this chapter we will examine aspects of object perception before considering the more complex issue of perceiving pictures.

Perceiving objects

In Chapter 1 we introduced the example of the dog guiding a blind human across the road. The guide dog must perceive its environment geocentrically in order to behave in the manner it does; it must respond to the edge of the pavement, the width of the road, and to the approaching vehicles in much the same way as a sighted human would. Both guide dogs and sighted humans respond to objects in terms of their locations and dimensions in three-dimensional space, rather than their locations and dimensions projected on the retina. That is, objects are normally seen as having constant dimensions despite changes in the retinal projections of their sizes and shapes. This is called perceptual constancy, and it is the prerequisite of any more complex perception like recognition.

In the preceding chapters we have described how geocentric perception of the environment is obtained from the patterns of stimulation reaching the eyes. An important feature of this approach has been the emphasis on a moving, active observer. By moving our eyes and our heads we gain information about new aspects of the world that would otherwise not be available. This information has to be extracted from the complex changes in stimulation that are caused by our own movements. As a result we obtain a geocentric representation of the direction, distance, and movement of objects with respect to the environment. Without this, we, and other animals, would be unable to carry out the co-ordinated activities necessary for survival. In addition, a geocentric representation provides the basis

for recognition, because it also conveys the defining properties of objects, like their orientation, location, size, and shape.

Perceiving object properties

Perception of object characteristics needs to conform to the frame of reference within which behaviour takes place, which will normally be the surface of the earth. We refer to such perceptions as geocentric. An active perceiver, whether human or animal, must recover the geocentric properties of objects, despite self motion. In the previous two chapters, we have described the frames of reference relative to which information can be represented, and from which geocentric information can be derived. The first, retinocentric, level is the one that is studied by neurophysiologists: information is coded in terms of the coordinate system of the retina. Retinocentric information describes image characteristics with respect to retinal coordinates, but it is not necessarily localised in the retina. The fact that a response occurs in the visual cortex, or elsewhere, need not preclude it from being retinocentric in character. The term retinocentric expresses the nature of the information, and not its anatomical site. As was discussed in Chapter 6, information about the location of objects can only be expressed in terms of a particular frame of reference. The retina generates neural signals that carry information about the location of stimulation with respect to a retinal frame of reference. Such information is then potentially available at any later stage in the visual system, and remains retinocentric unless transformed into another frame of reference.

This retinocentric level is essential for vision but it cannot serve as a basis for object perception or recognition because we have two eyes and both of them move. The next frame of reference uses an integrated binocular signal from the two eyes together with their movements to provide an egocentric frame of reference, the coordinate origin of which lies between the eyes. We perceive the directions of objects with respect to the egocentre. Visual direction alone would not enable us to locate objects in space; in order to do this we also need to determine the distance objects are away from us along the line of visual direction. There are many sources of information for the relative distances between objects – which is nearer or further – but these need to be anchored by information for object distance before objects can be adequately located in space. If information for the observer's own movements is incorporated, then objects can be represented in the geocentric frame of reference. When we have derived a geocentric representation of objects, we are in a position to behave appropriately with respect to them. The operation of these frames of reference applies equally to object characteristics such as motion, size, shape and orientation.

Perceptual constancies

In this section we will examine the phenomena known as the perceptual constancies, together with some of the explanations that have been proposed to

account for them. The term perceptual constancy refers to the fact that we normally perceive objects as having constant characteristics, like orientation, size, shape or colour, even when there are changes in the information about these characteristics that reaches the eye. However, the issue is really broader than is implied by this definition. A fundamental requirement for perception is that both the changing and constant characteristics of the environment should be accurately represented. This has to be accomplished despite changes in the pattern of stimulation reaching the eye due to an observer's own activities. We can perceive that objects with a constant physical size are not changing in size when we move towards or away from them or they move towards or away from us, and we can perceive an object's shape correctly even if it is seen from different directions. Without such abilities, recognition and identification would be impossible, as there would be no consistent description of an object to remember and make use of on subsequent occasions. It is equally necessary to perceive changes when they do occur. An object may move between locations in the environment, or undergo changes in orientation, but it must still be possible to recognise it correctly. Perhaps the most characteristic property of living things is that they can change both their shape and their location. A tree may be blown by the wind and alter its shape considerably. Animals can move about in the environment, and adopt a wide variety of postures. Biological shapes are highly variable, but our capacity for recognition is most striking for just these sorts of patterns, like those which define an individual's identity. Thus, an account of how we are able to perceive objects as constant when they are constant must explain how we can perceive changes in objects as well.

Orientation constancy

Orientation constancy is such a fundamental part of our experience that we rarely consider it as a remarkable achievement. Despite the many changes in the tilt of our eyes, head or body we perceive the world as upright. All you have to do to check this is to tilt your head to the side and view the world. Although the images of objects projected to your retinas are tilted, the world appears upright. In the terms described in Chapter 6 and illustrated in Figure 6.4, orientations are perceived geocentrically. The constraints of gravity affect all that we see and also our own posture. Receptors in the inner ear detect deviations of the head from gravitational vertical and send signals to the muscles to make rapid compensations so that we do not lose our balance; they are also involved in compensating for inclinations on the retina. The buildings that we see and the contents within them comply with the demands of gravity. It is only in fairgrounds and at science centres that tilted structures are intentionally made with the purpose of inducing disorientation in those who enter them (see Figures 6.7 and 7.5). The few people who have travelled beyond the earth's gravitational field have experienced a different visual world – one defined by patterncentric references alone. The internal structure of the space craft itself defines the frame of reference relative to which other objects are orientated and also relative to which the body itself is adjusted. For the rest of us who remain in closer contact to the earth's surface, we make continual

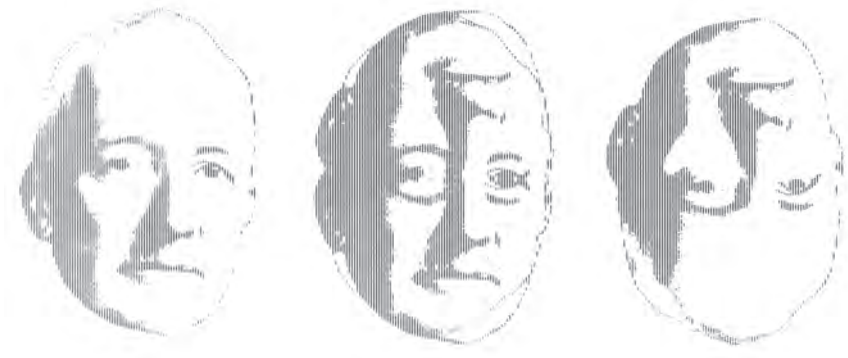


Figure 8.1 **An upright man** An upright, inverted, and combined view of George Malcolm Stratton (1865–1957).

compensations for variations in eye (and head) orientation so that objects appear to retain the same orientation. These everyday aspects of orientation constancy receive confirmation from laboratory studies.

In orientation constancy we compensate for changes in our orientation with respect to gravity so that objects appear to retain the same orientation despite their changing retinal orientation. This can be demonstrated under the reduced conditions of the laboratory: when only a single line of light is visible we can judge its orientation quite accurately even when the head or the whole body is tilted to the side. This is possible because of information provided by the otolith organs in the vestibular system; they signal the orientation of the head with respect to gravity and these signals are used to modify the information for orientation relative to the retina. The descriptions of the frames of reference involved in orientation constancy can be read in Chapter 6.

The normal relationship between objects and their projected images can be disrupted in several ways and one of the most noted early experiments in psychology was concerned with perceived orientation. The experiment was conducted by George Malcolm Stratton (Figure 8.1) at the end of the nineteenth century. He viewed the world with an upright retinal image (optically inverting the inversion in the eye) and described his experiences over a period of days. He wore lenses in front of his right eye, with the left occluded, for a period of three days. At the time he was one of Wundt's students at Leipzig, and made the initial announcement of his experiment at the Third International Congress of Psychology, held at Munich in 1896. On his return to California he extended the experiment by wearing the lenses for eight days. Upright vision with an inverted retinal image was then typically accounted for by either projection or eye movement theories. Stratton posed the question: "*Is* the inverted image a necessary condition of our seeing things in an upright position? The method of approaching the problem was to substitute an upright retinal image for the normal inverted one and watch the result". He was the watcher. At first, objects appeared

inverted, perceptual-motor coordination was disrupted and there was a distinction between where objects were seen and where they were thought to be. "By the third day things had thus been interconnected into a whole by piecing together the parts of the ever-changing visual fields." He did not report any aftereffects when the lenses were removed after three days, and he concluded that "the difficulty of seeing things upright by means of upright retinal images seems to consist solely in the resistance offered by long-established previous experience".

The first experiment was conducted entirely indoors, but when he repeated it over eight days in California he ventured outdoors, too. The pattern of experiences was similar for the first three days, and the adaptation became more established thereafter. There were, however, aftereffects when the lenses were removed: "On opening my eyes, the scene had a strange familiarity. The visual arrangement was immediately recognised as the old one of pre-experimental days; yet the reversal of everything from the order to which I had grown accustomed during the past week, gave the scene a surprising, bewildering air which lasted several hours. It was hardly the feeling, though, that things were upside down". The theoretical lesson to be drawn from the longer study was essentially the same as that from the shorter one: "harmony comes only after a tedious course of adjustment to the new conditions, and that the visual system has to build anew, growing from an isolated group of perceptions". The behavioural effects of adaptation to optical inversion have generally been confirmed in subsequent investigations, but the enigma of whether the world subsequently looks upright has proved harder to resolve. Ivo Kohler (1915–1985) wore a variety of distorting optical devices for long periods and found that both perception and action adapted to the new spatial relations. In one study he wore a binocular mirror-device (which inverted the images on both eyes) continually for almost four months. Initially everything appeared inverted and all actions based on visual input were misguided. After a period of days and weeks his behaviour was no longer disrupted by the inverting mirrors: he was able to reach for objects appropriately and even carry out complex skills such as skiing. Throughout the weeks of adaptation he kept detailed protocols and records of his performance on perceptual-motor tasks. When the mirrors were removed long-lasting aftereffects were reported: the world was transformed once more, but recovery was quicker than adaptation had been. "This experiment, the first of such long duration, was significant. In the first place, the aftereffects obtained were of optimal strength. In the second place, it gave rise to a number of peculiar aftereffects which I have already referred to as 'situational.' Not only curvatures, distortions, deviations, apparent movements, etc., were found to leave traces in the sensorium, but also the variations in intensity of these disturbances." Kohler's methods were in the phenomenological tradition of Gestalt psychology: "although my attitude toward this is somewhat ambivalent. I do not deny what is called 'organization' in perception, but I believe we have found a way to change this organization by prolonged exposure to certain kinds of stimulation." This ambiguity reflects an unresolved riddle associated with phenomenological reports of vision during inversion – whether the visual world appears normal after perceptual-motor adaptation. A factor that has rarely been taken into account is

Figure 8.2 Upright and inverted illustrations from around 1700 by an Italian artist called Giuseppe. The four upright engravings on the left are inverted on the right. By this pictorial device, even the gender and species boundaries can be crossed!

the restricted field of vision imposed by the optical devices. This results in visual conflict between the contents of the visual field and the border within which they occur. If the head moves upwards then the consequences for the visible frame (which is attached to the head) are the opposite of the visual field that can be seen within it. Moreover, inverted vision, using such restricted fields, has not been compared with normal vision under similar visual restrictions.

Measures of cortical activity (using fMRI) have failed to show any changes with optical inversion. One important lesson to be learned from Stratton's studies is the intricacy of the interactions between the visual and motor systems. Another is that the repetition in books and lectures of a 'fact' doesn't make it true! It is very widely believed that Stratton and others invariably found that the world appeared inverted when the optical apparatus was removed after adaptation; as so often, the reality is less conveniently clear-cut.

It is evident from Figure 8.1 that a picture of an inverted face is more difficult to recognise than that of an upright face, and this aspect of Stratton's studies led to further enquiries about recognition and orientation which were pursued by Wolfgang Köhler (see Figure 3.8). When discussing the optical inversion experiments, Köhler described the effect clearly: "For this experiment I select a picture, or outline-drawing of an object, which shows a conspicuous change in appearance when it is upside down. This is the case, for instance, with photographs of known or unknown persons. They change so much that what we call facial expression disappears almost entirely in the abnormal orientation". This observation has led to a veritable torrent of research papers in recent years, all of which confirm the difficulty of dealing with inverted and inclined faces. Artists have manipulated this paradox of orientation perception with skill for almost two thousand years and some examples are shown in Figure 8.2.

Similarly, written script, particularly if it is by hand, is difficult to read when presented upside down.

Location and motion constancy

One fundamental object property is its location, whether this is changing (i.e. the object is moving) or fixed. If behaviour is to be appropriately directed in the environment, then it must be guided by veridical information about the location of objects within some stable frame of reference. If the object is moving, then its velocity will need to be recovered. As we have seen in earlier chapters, this presents problems for biological visual systems, since the motion of an object's image on a retina may be caused by object movement, movements of the eye in the head, movements of the head on the body or movements of the whole body through the environment.

Consideration of retinal stimulation alone cannot account for geocentric visual perception. Many other sensory systems are involved in the recovery of the location and movement of objects in space. For example, information from the vestibular system is used to determine accelerations of the observer's head in three dimensions, and this is essential if the pattern of retinal change arising from a head movement is to be interpreted unambiguously. The integration of these sensory systems constitutes visual perception, and it would be wrong to suppose that visual perception can be understood solely in terms of light reaching the eye. Initial registration of retinocentric location and motion in each eye gives rise to a single binocular (cyclopean) retinocentric signal. This expresses location and motion with respect to a point located between the eyes. Since the two eyes move by equal amounts when version occurs, a single signal for eye movement can be combined with the binocular retinocentric signal to give an egocentric representation. This carries information for changes in angular direction with respect to the observer, despite movements of the eyes. If the observer's head moves (either with or independently of the rest of the body), then the eyes necessarily move too. This causes displacement of the retinal image, in a way that cannot be interpreted unless the extent of self motion and the distance of the object are known. Scaling the egocentric information by perceived distance gives the three-dimensional location of objects with respect to the observer. This can be corrected for the effects of self movement, since both are expressed in terms of three-dimensional displacements, to produce a geocentric representation of the position and movement of objects in the environment. Information for eye movement, self movement and distance may be derived from a number of sources, including visual ones. For example, when there are several objects at various distances, their relative displacement during self movement may provide information for their relative distance. At this level of analysis, the nature of the information required is defined, rather than its sources. The concern is with the rules and relationships that permit a geocentric representation of the external world by observers who can move in the environment, and whose sense organs can move with respect to the body. A geocentric representation is required for

effective action in the environment, and if a geocentric representation is stored in memory it can be matched to the same object when encountered on a subsequent occasion. There may of course be errors in the process of recovering a geocentric description. Eye movements, self movements or distance may not be accurately detected, and image displacements on the retina may be more or less than the signal to which they give rise. If so, then there will be errors in the perception of shape, size or movement, and objects may not be recognised.

Size constancy

Size constancy is easily observed, although we seldom notice that it takes place. In order to see this for yourself, hold up one hand with your palm facing you about 25 cm from your face, and the other with your arm extended. Separate your arms by an angle of about 30 deg, so that you need to turn your head to look from one hand to the other. When you do this, do your hands appear different in size? It is likely that they do not, in which case you are correctly perceiving their constant size despite large differences in the angles they subtend at your eyes. How can size constancy be explained? One possibility is that you see your hands as a constant size because they are familiar objects, and you already know from experience what size they are. This view, which might be the explanation suggested by common sense, cannot be correct. It would mean that size constancy would only apply to objects which we recognise, and whose size we know about from past experience. Clearly it would also mean that recognition of objects precedes perception of their characteristics, and this presents logical problems. An alternative approach could be to argue that we are mistaken in supposing that constancy is a problem to be solved by the visual system.

Suppose that an object like a car is seen at various distances. If only the optical projection of the car is considered, then our ability to see it as a constant size seems in need of explanation. The angle subtended at the eye by the car is apparently ambiguous, since the same angular subtense could represent a wide range of different-sized vehicles, at different distances. However in practice we would not rely only on the visual angle of an object in order to determine its size. Other sources of information may be available in the retinal image that remove the ambiguity, and determine the perceived size. One suggestion made by Gibson was that the amount of adjacent background texture obscured by an object remains constant, despite changes in its distance and angular size. This would then constitute an invariant property of the optic array, which directly specifies size. While there is relatively little information available about which optical invariants are significant, and how they may be detected, the general point made by Gibson about the lack of real ambiguity in the retinal projection is a very important one.

Size-distance invariance

There are other sources of information that determine an object's apparent size, as well as purely optical invariants, and the most important of these is perceived

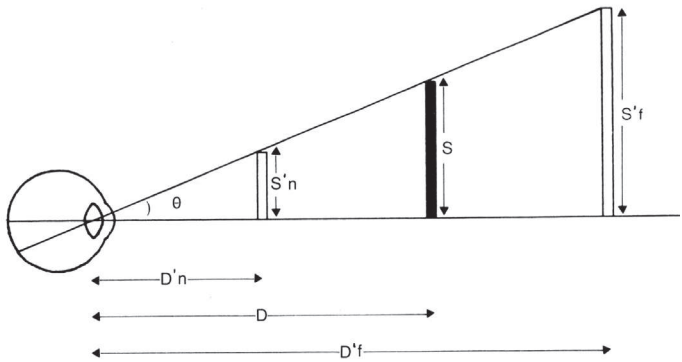


Figure 8.3 The relationship between perceived size, perceived distance and visual angle, as expressed by size-distance invariance. An object of a given physical size (S), at a given physical distance (D), subtends an angle of θ deg at the eye. Such a visual angle could correspond to an infinite number of different-sized objects at different distances, all of which have the same ratio of size to distance, and therefore the same visual angle. By size-distance invariance, the perceived size of an object (S′) depends on its visual angle and its perceived distance (D′). Thus, if the object is incorrectly seen as nearer than D (D′n), then it will appear to have a size of S′n; similarly with too large a perceived distance (D′f), the perceived size is S′f, so that the ratio of S′ to D′′ is constant. In general, S′ = (D′.S)/D, or S′ = D′.tan θ, since tan θ = S/D.

object distance. Cues to object distance were discussed in Chapter 6, and there is good empirical evidence that perceived size is linked to perceived distance.

The relationship is known as size-distance invariance, and it can be expressed by the statement that perceived size is proportional to the product of visual angle and perceived distance; or, equivalently, that the ratio of perceived size to perceived distance is proportional to the visual angle. The geometry of this relationship is illustrated in Figure 8.3. Since visual angle is the ratio of physical size to physical distance, another form of the equation is that perceived size is proportional to the product of physical size and the ratio of perceived distance to physical distance. Therefore, if perceived and physical distance are the same, their ratio is one, and perceived size will be constant if physical size is constant, even though distance varies. For an object of constant physical size, the product of visual angle and perceived distance is a constant (invariant) characteristic, despite changes in distance. If angular size increases and perceived distance decreases, as would happen with an approaching object, then perceived size will remain constant.

Size constancy can be regarded as a special case of size-distance invariance, which can also account for situations where errors in perceived size occur. Suppose that the perceived distance of an object is increased by a change in the cues available for perceived distance, such as a decrease in the convergence of the eyes. Its visual angle remains the same, because its physical distance has not altered, so the product of visual angle and perceived distance will be larger. That is, the perceived size of an object with a constant visual angle should increase

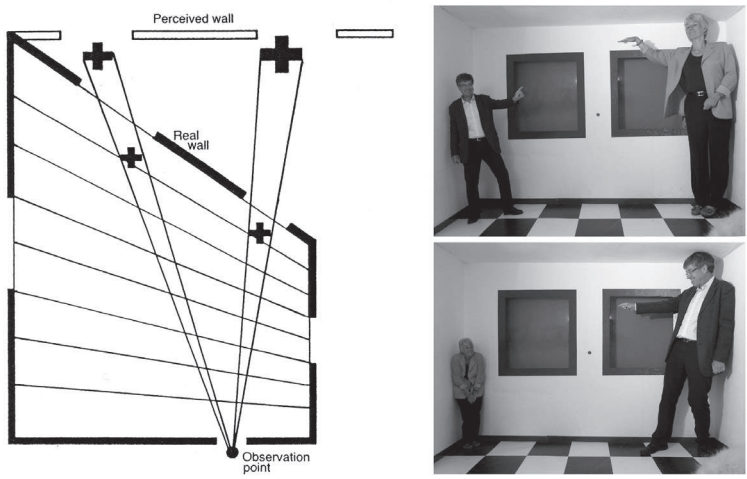


Figure 8.4 Left, the layout of an Ames room seen from above. The actual shape of the room is trapezoidal, with the far wall sloping at an angle towards the observer. However all the available cues from linear perspective, texture gradients, etc. are adjusted to indicate that the far wall is at right angles to the observer. For example, window frames are trapezoidal, but give rectangular projections at the eye due to the slope of the wall in which they are set. The far wall is therefore seen in the position shown, and as a result objects placed in the room appear to be distorted in size. Equal-sized objects appear much larger when seen on the right than on the left. These perceived sizes are what would be expected from size-distance invariance, given the errors in perceived distance. Right, photographs of the same two individuals, of about the same height, at either side of an Ames room.

if its perceived distance increases, and decrease if perceived distance decreases. This outcome can readily be observed and is termed Emmert's law, although it has been reported several times in recent centuries. An afterimage can be produced by fixating, for a minute or so, on the white dot in the centre of Figure 7.10. The afterimage of the grid pattern can then be observed against a close surface like your hand, or a further one, such as the ceiling. You should find that the apparent size of the afterimage changes, so as to appear larger when seen against the further surface. Clearly, there must be good cues to the distance of the surfaces. The retinal area stimulated necessarily remains a constant size, and is seen at the apparent distance of the background as a result of the equidistance tendency.

The Ames room is named after the American ophthalmologist Adelbert Ames II (1880–1955) and it demonstrates size-distance invariance very clearly. Figure 8.4 left illustrates the geometry of an Ames room. Although the far wall is at an oblique angle to the observer, all the internal features, such as the skirting boards and window frames, have been shaped to appear as if on a front-parallel surface. With monocular observation, the effective cues to distance (principally linear perspective) correspond to a rectangular room, and the perceived distance of the far wall differs from its physical distance. As a result there are remarkable

effects on the perceived size of objects placed within the room. Two similarly sized people standing in each corner look to be very different in size, as is shown in Figure 8.4 right. If the people change places, they seem to grow and shrink. It is not necessary to suppose that this has anything to do with knowing that rooms are generally rectangular, because the effects are exactly as would be predicted on the basis of size-distance invariance, given the available cues to distance. The apparent sizes of objects in the room are appropriate to the product of their visual angles (which depend on the physical size and the physical distance) and the perceived distance. Perhaps the most striking aspect of the Ames room is that knowledge about the objects within it has little or no effect on perception. Even if people are observed moving around in the room, they still appear to shrink and expand. It is often stated that if these people are of special significance to the observer, they may be seen veridically and the room as non-rectangular. However, following initial reports, there has been little evidence to confirm this influence on perception of knowledge about the actual sizes of such familiar objects.

The account of geocentric perception given above was based on a similar principle to size-distance invariance. The scaling of changes in egocentric direction by perceived distance is equivalent to the scaling of visual angle by perceived distance in size-distance invariance. The equivalence is apparent if you think of a difference in the egocentric direction between the edges of an object as defining visual angle, and a difference in egocentric direction of a single point over time as defining movement. Our account of geocentric perception predicts both size constancy and location constancy; that is, our ability to see stationary objects as stationary despite observer movement. Location constancy is a special case of a more general process of geocentric motion perception. An adequate explanation of spatial perception will account for both location and size constancies, as well as the systematic ways in which they break down.

Shape constancy

The same principles can be extended to explain shape constancy. Consider a rectangular object like this book. When closed and face-on, the retinal projection of the book is rectangular, and corresponds to the physical shape. If the book is rotated about a vertical axis, the retinal projection of the cover is altered to a trapezoid, because the far edge now subtends a smaller visual angle than the near edge. The principle of size-distance invariance is sufficient to account for our ability to continue to see the cover as rectangular even when it has been rotated. If there are adequate cues to distance, the perceived size of each edge will correspond to its physical size, and thus the overall perceived shape will correspond to the physical shape.

An alternative view of how we achieve perceptual constancy suggests that it is based on assumptions about the structure of physical objects. The best known is called the rigidity assumption, because it consists of interpreting retinal changes in terms of the rotation and translation of rigid three-dimensional objects. For example, the changes in the contours of a book as it rotates about a vertical axis

could be produced by a large set of environmental events, in which the book itself undergoes changes in physical size and shape. If it is assumed in advance that objects are rigid and unable to change shape, then the retinal transformations allow only one interpretation. The difficulty with the rigidity assumption, and other sorts of a priori principles in perception, is that it constrains what can be seen. Objects can in fact change shape, since some are elastic. We can evidently see non-rigid transformations when they occur. A rigidity assumption may not be necessary if sufficient information is available about an object, in addition to its retinal transformations.

Underlying much discussion of perceptual constancy is the conception of a retinal image which is intrinsically ambiguous. The same optical projection on the retina can arise from many different physical structures, and yet our perception normally corresponds to the one which is actually present. This point of view has led to the development of various explanations of how we may resolve the ambiguity. It is unarguable that the retinal image of a given object considered in isolation provides ambiguous information for its size, location and shape. Like any ambiguity, this can be resolved if other information is available. The major theoretical approaches to this have been to incorporate past knowledge, prior assumptions, other concurrent sources of visual information such as distance, or other information in the retinal image like texture and perspective. However, objects are seldom seen in isolation outside the laboratory, and if they are, they may well be perceived incorrectly. The appearance of an object can be changed by the visual background against which it is set, and there is little doubt that the visual system interprets the retinal image of an object in the context of other retinal patterns, although we know little about how this comes about. As has been demonstrated, information about object distance is critical in determining perceived size, motion and shape. These perceptual characteristics can best be understood as aspects of spatial perception, which must be geocentric if behaviour is to be appropriately directed. In addition, a geocentric representation provides a description of objects in terms which would permit subsequent recognition.

Colour and lightness constancy

Two further perceptual constancies may be less closely related to the processes of spatial perception. Colour constancy refers to the fact that we see objects as being a constant colour even when the wavelength of ambient illumination changes. Under white light, an object will reflect a part of the total spectrum, which determines its colour. If the ambient illumination is restricted to, say, the longer wavelengths (so as to appear red), the object will continue to appear the same colour as before, even though it now reflects a quite different range of wavelengths. In effect we are able to compensate an object's apparent colour for any bias in the colour of the prevailing illumination. In order for this to happen, there must be sufficient information about ambient illumination. Thus, if an object is illuminated by a narrowly focused source of coloured light, which does not impinge on the rest of the scene, then colour constancy will fail, and the object

will seem to be a colour determined by the fraction of the incident wavelengths which it reflects.

A similar process, known as lightness constancy, occurs with changes in the intensity of ambient illumination. A piece of coal will appear black and a sheet of white paper will appear white, even when the intensity of the light they reflect is changed over a wide range. Even more strikingly, if the paper is in shadow, it may reflect less light than the coal, but the coal will still appear darker than the paper. Again, this is dependent on being able to perceive the intensity of the ambient illumination. If the paper and the coal are placed side-by-side and the coal is illuminated by a hidden spotlight, it can be made to appear whiter than the paper. These constancies are also influenced by perception of the three-dimensional space in which illuminated objects are placed. It has been found that the perceived lightness of a surface was altered if its apparent distance (and thus the shape of the object of which it was a part) was perceived to change.

Recognising objects

All animals need to respond appropriately to aspects of their environment that are important for their survival – for their sustenance, shelter and reproduction. Such aspects are three-dimensional, and so the general rule advanced above is that perception is geocentric. This applies to guide dogs before their specialised training. They can avoid objects, bite them, paw them or climb onto them without any training by humans. Guide dogs require extensive training before they are assigned to a blind owner because they need to behave unnaturally. They are trained to stop at small steps that they could easily negotiate and to avoid projecting objects that they could easily run under. In effect, they are trained to respond as if the space they occupy is equivalent to the dimensions of a human. It is rather like starting to drive a car, when we gradually learn to control a vehicle that occupies a greater volume of space than we do. Guide dogs are trained to discriminate features of their environment that are not intrinsically important to their behaviour. For example, a dog can be trained to discriminate between steps that are too low to interfere with the blind person's gait, and steps that are higher and might be tripped over. The dog is trained to respond differentially to two different states of the world (not to two different retinal projections), and so there will be some stored internal representation of these against which any present state can be compared. That is, the dog can recognise the distinction between the heights of steps. This behaviour is not restricted to the particular steps for which training occurred, as the guide dog would be expected to respond appropriately to examples of steps not previously encountered. Therefore, recognition involves two complementary aspects – discrimination and generalisation.

Discrimination and generalisation

Discrimination concerns assigning different behaviours to specific objects or object properties, like two different heights of step. In addition, these discriminatory

behaviours can be generalised to instances of the same properties in other objects, as when a guide dog responds correctly to a different road from those on which it was trained. The points about discrimination and generalisation mentioned above can be illustrated further with an example more readily to hand. If the copy of this book that you are reading were to be placed with a number of different books, you could no doubt recognise it correctly. As was discussed above in connection with perceptual constancies, you would see the shapes, sizes and colours of the books correctly, even if they were at different distances, oriented at random with respect to the surface on which they were placed and illuminated with different intensities or wavelengths of light. You would therefore have a perceptual representation of each book that could be compared with a stored representation of previous perception. You could of course recognise this particular book on the basis of the text on the cover, but even if any words were hidden you would probably still be successful. The information you would use would relate to the size, shape and possibly colour of the book. Evidently we must have this sort of information available in order to discriminate. If another copy of the same book were put in the place of this one, the task would become much more difficult. You might well suppose that the replacement copy was the original, and incorrectly recognise it. This would be an instance of generalisation, in which there is sufficient similarity between the characteristics of two objects for the response to both to be the same. In order to pick out this particular copy, it would be necessary to store some information about it that was sufficiently specific to allow discrimination between copies, like marks or damage to the cover. Generalisation gives an indication of the nature of the representation that has been stored, by comparing the characteristics of objects that are confused. It demonstrates the occurrence of recognition, and shows the basis on which this has taken place. For instance, you would be likely to recognise this book if it were upside down. Your response would generalise to different orientations of the same object, which means that the stored representation does not include the particular orientation in which the book was first seen. None the less, you would also perceive that its orientation had changed.

Both generalisation and discrimination are indications that recognition has taken place. In studying recognition in animals and human infants, behavioural measures of recognition are necessary. With normal adult humans, recognition is more usually assessed by the seemingly simpler method of asking an observer for a verbal response. While this has been the basis of much research on recognition, the issues addressed by this research often concern the linguistic rather than the perceptual representation of information about the world.

The investigation of recognition and identification has a very long history. The early philosophers of China and Greece confronted the issue of how we can recognise an object as a member of a general class, and also as a particular instance of that class. While philosophers concentrated their thoughts on the problems of knowing about real objects, psychologists have tended to study recognition in terms of simplified laboratory tasks, generally involving two-dimensional pictorial patterns. Would you mistake a picture of this book for the book itself?

Would you expect a guide dog to do so? Much research on recognition has been based on supposing that this would be the case. Since pictorial images are so widespread in both research and everyday life, it is important to know how the recognition processes that apply to real objects may generalise to pictures.

Many animals, besides humans, have well-developed capacities for discrimination and generalisation. These abilities have been the subject of extensive investigations, usually in the context of studying the mechanisms of learning. While a full understanding of recognition would involve consideration of the processes responsible for learning, the initial step is the acquisition of perceptual information necessary for recognition. The processes responsible for perceptual constancies provide a geocentric description of objects which capture their unchanging, intrinsic characteristics. These processes also convey information about changes in objects that do not alter their identity, such as their location and orientation, and the representation they provide is therefore a suitable basis for recognition.

Pictures

In the case of humans, significant activities are not necessarily restricted to the primary demands of survival and reproduction. We may seek a variety of sources of visual stimulation, many of which are nowadays cultural products. Some of these, like painted pictures, may be almost as old as the human species itself. Others, such as photographic or digitised images, are more recent. Writing and printed text are visual stimuli of particular complexity, because the shapes bear little visual relationship to the objects or meanings they convey. Thus, human survival seems to depend on intellectual as well as nutritional sustenance, and in the case of the former pictures appear to satisfy a universal appetite. The use of created patterns of visual stimulation as representations of other things is one of the most characteristic features of modern culture. Of course, pictures may sometimes be abstract, and appreciated for the shapes they contain, but it is very hard not to attribute meaning into anything we look at. Pictures can be both objects in themselves (a surface with marks on it like a painting or computer screen) and representations of other objects. One consequence of this is that pictures are likely to be more complex stimuli to process than solid objects.

How might animals, such as our guide dog, respond to pictures? What does a dog see when it is looking at a television programme? Little experimental evidence exists, despite a plethora of anecdotes. We might expect a response on the basis of stimulus generalisation; that is, in terms of the extent to which a pictorial pattern has common features with the thing it represents. Certainly, pictorial stimuli are often used to study animal vision, but these pictures are usually quite abstract, and the animal learns to associate a particular visual feature (such as the number of dots or the orientation or colour of lines) with a particular response. It is much harder to know if an animal's perception of a pictorial representation in any way resembles ours. There has been much research on the possibility of animal language (a different type of representation of reality) and there are reports of

the production of pictures by higher primates and even elephants, but this does not help to answer the question. The issue is unresolved at present, but perhaps deserves future study.

Some aspects of picture perception and production were described in Chapter 2. Here we will focus on the perception by humans of pictures as representations, because we now derive so much of our information about the world through pictures. We will also discuss some of the difficulties of perceiving and working with these sorts of stimuli. In addition, we will consider the perception of pictorial representations that have some of the three-dimensional properties of real objects – once sometimes called 'virtual reality'. Although this term is now rather dated, the pursuit of realism in pictorial representation, especially cinematic, is undiminished. Computer-generated images (moving or stationary) can be almost indistinguishable from a conventional photographic image of a real scene, and the technological and economic drive is such that this convergence between the imagery of real and simulated scenes is only likely to accelerate. We begin by considering the principles underlying the relationship between pictures and the objects they represent.

Pictures and objects

Most experiments on object recognition use pictures of objects rather than the objects themselves. Accordingly, we know quite a lot about picture recognition. In order for this knowledge to be of use in furthering our understanding of object recognition we need to appreciate the relationship between pictures and objects. This was discussed in Chapter 2, particularly with respect to a painting by René Magritte (Figure 2.16), in which a curved pipe was pictured above text stating that it was not a pipe. Magritte was contrasting two different forms of representation – pictorial images and written text. Both are in fact pictorial representations, but one concerns the spatial aspects of the particular object and the other the conventional description of the class of objects – pipes. Magritte exploited this distinction with skill because the pictorial representation of the curved briar pipe was readily recognisable. The orientation he chose for its representation was not arbitrary; had he painted it from other orientations its recognition would not have been as rapid. For example, the four pictures shown in Figure 8.5 are photographs of the same pipe as was used for producing Figure 2.16, but they were in different orientations with respect to the camera. Some might not even be recognised as representing a pipe if they were to be presented in isolation and without any context. You could try this out by showing someone the picture in the lower left quadrant on its own, by covering up the others, and asking them to say what object it would correspond to. We would not have the same difficulty in recognising an actual pipe in a number of orientations, because we would have information for its orientation in space and its dimensions. Therefore, if we are dealing with pictures of objects we need to consider the viewpoint from which the pictures are taken, and the orientation in which the picture itself is presented.

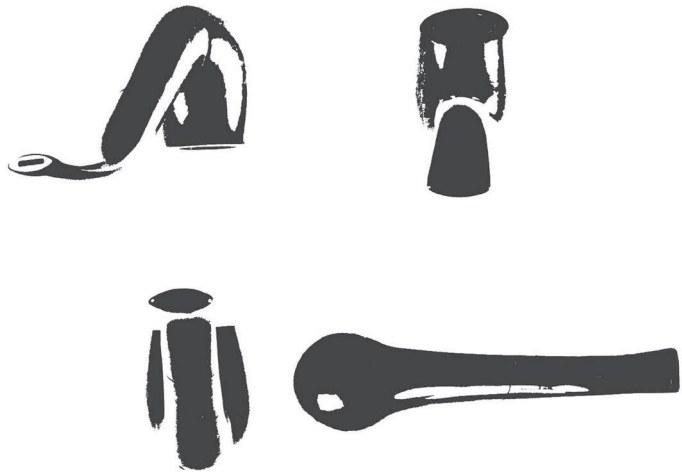


Figure 8.5 Four photographs of the same pipe as shown in Figure 2.16 but taken from different viewpoints.

Pictures and their perception have often proved a fascination for visual scientists. Pictures have an advantage over objects in so far as we can manipulate their characteristics much more readily than we can those of objects. Therefore experimental studies are much easier to conduct with pictures, although there are many reasons for concluding that perception of a picture of an object is not the same as perception of the object itself.

Orientation and bearing

It is clear from the example above that some pictures of the same object can be recognised more easily than others, according to the viewpoint from which they are imaged. Clearly this only applies to asymmetrical objects: different views of a sphere would be the same, but almost all objects in our environment are asymmetrical. Even a symmetrical object like a cube can be pictured in many different ways, depending upon the orientation of the faces with respect to the viewpoint. With a camera mounted on a tripod, imagine photographing a cube with its near face in the frontal plane and the centre aligned with the optical axis of the camera. The outline of the photographic image would be a square (see Figure 8.6a), and the same configuration would result from any rotation of the cube around the optical axis of the camera. Figure 8.6b shows another view, with the face of the cube rotated by 45 deg with respect to the camera. The only difference between these two pictures is their orientation, but this is a critical one for our perception. We can discriminate differences based upon orientation alone, so that even though we can recognise that the configuration is a square in both instances, we can also discriminate the difference between them. In fact, in this example we are likely to give different names to the same configuration, based solely on

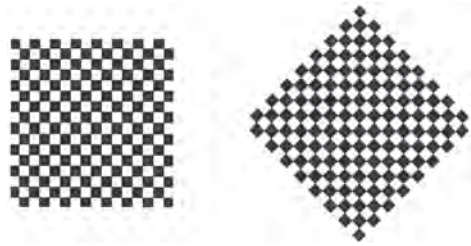


Figure 8.6 Photographs of a cube (a) with its sides vertical, and (b) with its sides inclined by 45 deg.

orientation: (a) would be called a square, whereas (b) would be described as a diamond.

We will discuss the perception of pictures of a cube in more detail, in order to make clear the basic principles of representation. Changes in the orientation of pictorial images can be achieved in two ways: the cube could be rotated with respect to the stationary camera, or the camera could be rotated with respect to the stationary cube. The resulting photographs would be indistinguishable if the rotations had been equivalent. This projective equivalence would also apply to an observer but the two states would not be perceptually equivalent. If a square is presented vertically and at 45 deg we can discriminate the difference with ease. On the other hand, if we view a square with the head vertical and then with the head rotated so that the square is at 45 deg with respect to the retina, the two will not be easy to discriminate because of orientation constancy. That is, we make compensations for the inclination of the head to gravity.

The photographs of a cube we have considered so far have shown only one of its faces, and its rotations have been around the optical axis of the camera. In order to picture two faces of the cube it is necessary to rotate it around a vertical or horizontal axis, and three faces are imaged with rotation about both (see Figure 8.7). When all three faces are pictured there is a wide range of viewpoints that can be adopted to display varying proportions of each. Only when a diagonal axis of the cube is in the optical axis of the camera will the area and configuration of the three imaged faces be equal.

Parallel sides on a face of a cube converge in the photographs shown in Figure 8.7. The degree of convergence depends upon the characteristics of the lens used in the camera, the size of the cube and the distance from the camera. The photographs of the same cube shown in Figure 8.8 were taken with different lenses, and they have been enlarged to the same pictorial dimensions. The focal length of the lens in a conventional 35 mm camera is about 50 mm; Figure 8.8a shows a photograph taken with such a lens – the parallel sides converge by a moderate amount, and much more than in the case of a long focal length (200 mm) telephoto lens (Figure 8.8b). The convergence is considerable with the wide angle lens used to produce Figure 8.8c. If the eye is to be likened to a camera, then we should try to compare like with like. The focal length of the eye varies from

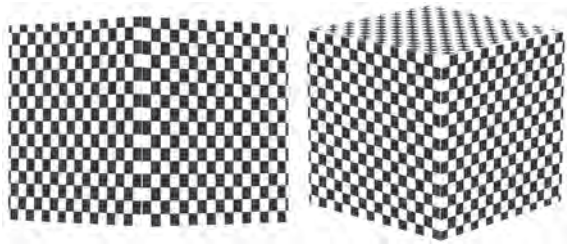


Figure 8.7 Two photographs of a cube (a) with two faces imaged, and (b) with three faces imaged.

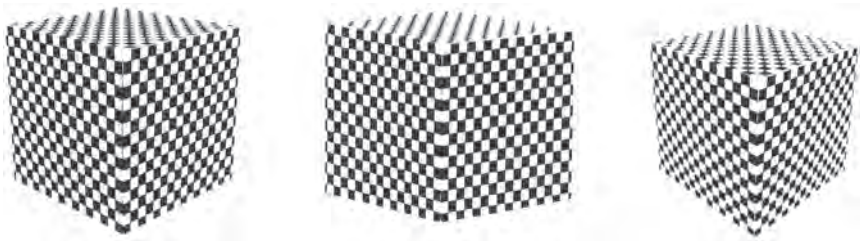


Figure 8.8 Three photographs of the same cube taken with (a) a 50 mm lens, (b) a 200 mm lens, and (c) a 20 mm lens.

around 14 to 17 mm, according to the state of accommodation. A lens with these characteristics in a camera would produce an image with massive distortions – straight lines would appear curved and the convergence of parallel lines would be considerable. Clearly, what we see does not correspond to the image formed in a camera with similar optical properties to the eye.

Stereotypical viewpoints

No photograph of a cube can be said to represent adequately a cube as an object, because they can, at best, only show three of the cube's six faces. The object pictured can only be a cube if it is assumed that the unseen faces correspond to those that are imaged. Exactly the same set of photographs to those shown above could have been produced with an object having only three connected faces rather than six – it could have been an empty shell. Despite this shortcoming we are remarkably good at recognising pictures of objects as representations of the objects. However, as was hinted above, not all pictures of the same object are treated equivalently. Returning to Magritte's pictured pipe, some viewpoints are more readily recognised than others. These have been called canonical, typical, or stereotypical views of an object, and they apply principally to asymmetrical objects like pipes and people. What are the features that the stereotypical view has, that render it more readily recognisable? Magritte's pipe provides a good vehicle for considering this issue. All the photographs of pipes shown in Figure

Figure 8.9 Silhouettes of a curved briar pipe viewed from (a) the side, (b) the front/back, and (c) the top/bottom.

8.5 are in accurate central perspective, as is that in Figure 2.16. Therefore, accurate representation is not the feature that distinguishes the stereotypical view from others. Many objects have a fixed polarity; they have a defined orientation with respect to the surface of the earth. This applies to people, who maintain a fixed posture with respect to gravity, and it applies to pipes, because tobacco will be retained in the bowl more readily in a particular orientation. Therefore, the stereotypical view needs to retain the normal orientation of the object. The normal orientation is retained in several of the photographs in Figure 8.5, but they are not stereotypical views because they foreshorten a critical dimension of the pipe – that along the stem.

The structure of any object can be described in terms of a set of three-dimensional coordinates of all points on its surface. This set of coordinates would enable its surface structure to be reproduced. Another way of achieving such a structural description would be to take three projections of the object from three orthogonal directions. This is rather like taking a silhouette of an object from three viewpoints at right angles to one another, and this has been done for a pipe in Figure 8.9. One silhouette corresponds to a view from the side, another from the front, and a third from above. (Note that this is treating the pipe as though it has a clearly defined side, front and top.) Each silhouette maximally foreshortens one of the dimensions of the pipe. Of the three silhouettes one is more readily recognisable than the other two as representing a pipe, and that is the side view (Figure 8.9a). The most difficult silhouette to recognise is that from the front (Figure 8.9b); in fact there is nothing in the silhouette to distinguish between a view from the front or the back. The view from above or below (Figure 8.9c) does not distinguish between a curved and a straight-stemmed pipe. Accordingly, the stereotypical view corresponds to the most recognisable silhouette of an object. It also corresponds to the viewpoint that minimally foreshortens the most asymmetrical dimension of the object. It will be noted that in Figure 8.5, as in Magritte's painting, the viewpoint was shifted slightly from the side view, so that the circularity of the bowl could be represented as an ellipse. Thus the three dimensions of the pipe are represented in the picture, but the most asymmetrical dimension is least foreshortened.

This analysis is easier with pipes than with people, because the surfaces are generally smoother with pipes. Nonetheless, the same general principles apply: silhouettes of heads would show the same advantage for the profile view over

Figure 8.10 Three photographs of a head viewed from (a) the side, (b) the front, and (c) the top.

that from front/back or from above (it would be difficult to represent the head from below!). There is a great deal of significance attached to the patterning of features on people's heads, and so rather than take silhouettes we could take normal photographs from three orthogonal viewpoints. If we consider only the head, then it could be photographed from the front, the side and from above, as shown in Figure 8.10. If these photographs were reduced to silhouettes then the profile would be recognisable but the other two would not. When the facial details are visible then it is not so obvious from which viewpoint the face can be recognised most easily. It is clear which is the poorest – the view from above (Figure 8.10c), but the other two seem equally informative. Thus, when the surface detail can be seen, it seems as though there are two views of the human head that are readily recognisable (Figures 8.10a and b). It is not coincidental that criminal 'mug shots' are taken from profile and full-face views, and that most portraits are in three-quarter profile, so that dimensions of the protruding facial features are not lost.

Photographs and drawings

Cameras are designed to produce pictures that are in central perspective; most of the pictures that we see in newspapers and on television have been derived from cameras and so are in perspective. The situation is somewhat different when we consider drawing and other graphic arts. Perspective arrived rather late on the scene in Western art, and cannot be considered in any way as universal. Paintings were made in caves from about 40,000 years ago, and we can recognise the objects (usually animals) that are depicted. They are nearly all represented in the stereotypical view – outline profiles of bison, deer and horses. Thus, from the earliest examples of art outlines have been used to describe and delineate representations of objects. When we draw objects we initially define their boundaries with lines, and often do not proceed any further. Outlines or contours are very informative, and they can be sufficient to establish the relationship between a drawing and the object it is intended to represent.

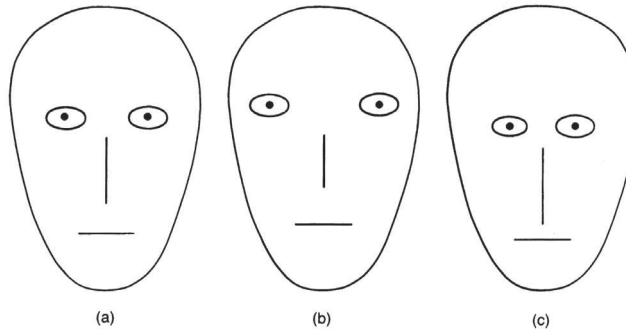


Figure 8.11 Examples of schematic faces. Brunswik produced nearly 200 schematic faces, all within the same oval outline, by manipulating: the height and separation of the eyes, the height and length of the nose, and the height of the mouth. Of the three examples shown here (a) corresponds to the neutral standard, (b) was rated by observers as young, good, likeable and beautiful, whereas (c) was rated as sad, old, ugly and unlikeable.

Drawings or outlines are used extensively in perceptual research, far more so than photographs. Outline drawings are used so widely because, in large measure, they are so easy to produce and manipulate. The situation has been compounded by the onset of interactive computer systems which enable an observer to manipulate characteristics of the outline display. This is one of the factors that has led to the burgeoning of research on visual illusions – they can readily be drawn and manipulated. Indeed, simple outline drawings have even been used to determine complex perceptual dimensions, like facial expressions, as in the experiments illustrated in Figure 8.11. There are consistencies in the emotions observers attach to these simple outlines, but it is difficult to know how to interpret them. In order to do so, we would need to know how the outlines relate to static views that are more representative of actual faces (e.g., photographs), and then how photographs relate to actual faces. Neither of these steps is understood at present.

Contours in drawings are abrupt changes in the amount of light reflected from a surface; they correspond to the boundaries of an object, but there are rarely such well-defined transitions between light and dark in the original scene. One of the problems that has beset workers in the area of computational vision is to define the boundaries of an object in a well-structured scene. Therefore, an outline drawing of an object is an abstraction that does not correspond readily to any features of the pattern of light projected from a scene. We are exposed to such simplified representations from a very early age, and so we learn how to interpret them. Experiments have established that we extract information not only from contours but also from the regions of a drawing where contours change direction or intersect.

It is clear that humans have produced pictures for much longer than recorded history, and a variety of styles can be discerned. Linear perspective is amongst the most contrived because it requires picture-makers to forgo their perception of space and to record visual angles from a single station point. Unlike the art

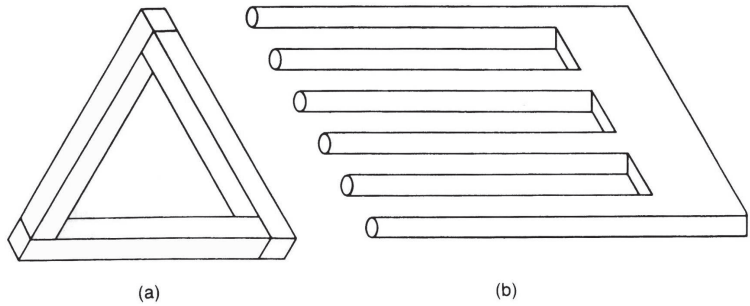


Figure 8.12 Impossible figures: (a) an impossible triangle, and (b) the devil's pitchfork.

forms that preceded it, linear perspective set out specific rules for representing the sizes of objects on the picture plane and, more importantly, the dimensions of texture on the receding ground plane. These were determined by the distance of the station point from the picture plane and the height of the station point from the ground plane (see Figure 2.17). Thus, all the 'painter's cues' to distance can be enlisted to increase the allusion to distance on the flat surface of the picture. Artists have to learn to draw in linear perspective and similarly observers have to learn to interpret pictures so produced. The drawing can be assisted by all manner of aids, the simplest of which were suggested by Alberti and Leonardo in the fifteenth century. If the picture plane is replaced by a pane of glass then, with a single eye at a fixed position (the station point), the contours in the scene can be traced on the glass surface: the tracing will be in accurate linear perspective.

The rules of linear perspective provide a consistent way of treating the dimension that the picture itself lacks – depth or distance. When three dimensions are reduced to two it is possible to play tricks with the transition, and to create worlds on paper that could not exist in three-dimensional space. Figure 8.12 illustrates two 'impossible' figures that employ perspective capriciously. The impossible triangle is in accurate perspective at all the corners, but the connecting lines create the impossibility. The devil's pitchfork is another example that is based upon the minimum requirements for representing a cylinder and a rectangular bar: the twelve horizontal lines are connected to six ellipses on the left, representing six cylinders, but they constitute only four rectangular bars on the right.

Allied to these impossible figures are others that are not sufficiently specified to be perceived unambiguously. Perhaps the most famous of these is called the Necker cube, after the Swiss crystallographer Louis Necker who first described it in 1832 (Figure 2.29c). He was observing a crystal under the microscope and noted that its structure appeared to alternate in depth. The Necker cube is an outline figure that is equivalent to a skeleton cube imaged with a telephoto lens so that the sides remain parallel (see Figure 8.13a). The figure appears to fluctuate in orientation, at one moment it seems to be pointing down and to the left, then it flips to appear pointing up and to the right. Note how the fluctuation is less pronounced for a representation of a skeleton cube which has converging

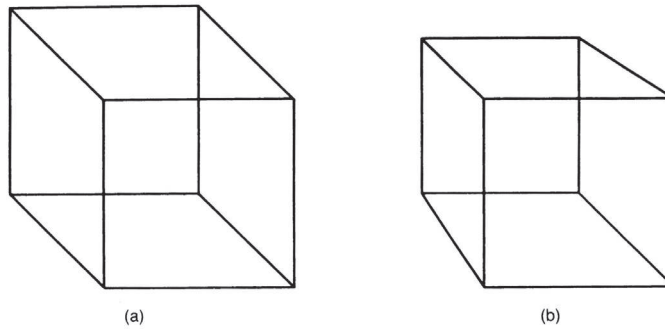


Figure 8.13 Necker cubes drawn with (a) parallel and (b) converging sides.

Figure 8.14 Photographic detail of a fifth or sixth century Roman mosaic in Domus dei Tappeti di Pietra, Ravenna, illustrating spatially ambiguous motifs; the eight radiating lozenges form sides of cubes, but a given lozenge could be the face of two cubes. The tesserae in the original are coloured black, off-white and light brown

sides (Figure 8.13b). Some artists have manipulated a related form of perceptual ambiguity in which the same outline can be seen as representing two different objects, only one of which is visible at once.

Similar patterns that produce alternating appearances, like those with the Necker cube, are of ancient origin. The mosaic shown in Figure 8.14 utilises the lozenge (diamond-shaped) motif: the lozenges radiate from a central point and vary in colour. They are symmetrical, having angles of 45 and 135 deg, so that eight lozenges radiate from a point to complete a circle. At the outer boundaries of the lozenges, squares are bordered by chevrons giving the impression of cubes with patterned faces. However, the cubes are spatially ambiguous because each lozenge can be a face of two cubes.

Perceiving pictures

We perceive pictures in the same way that we perceive other objects in the environment: they are flat, mostly on paper, but occasionally in frames. There is nothing peculiar about pictures as objects. However, this is not what is usually meant by picture perception. The term refers to the interpretation of the marks on the surface of the picture – to the recognition of the marks as referring to objects that are not actually there. Thus, pictures have a dual reality: they exist as objects and, in addition, the marks signify other objects and spaces than those they occupy. In the first sense, pictures pose the same problems as pipes and people; they have locations and dimensions that need to be resolved by the perceiver. In the second sense, they are paradoxical because they allude to another space or to other objects.

It is possible to arrange conditions such that a picture is not seen on the surface it occupies. Such works are called trompe l'oeil because they 'fool the eye'. Successful trompe l'oeil paintings are very rare, and they are exceedingly contrived. One example can be found in the National Gallery, London; it is a perspective cabinet painted by the seventeenth century Dutch artist, Samuel van Hoogstraten. The scene represented is an interior of a house with many interconnecting rooms, which can be seen through open doors. The interior of the cabinet can only be viewed through a peephole. What is of interest is that the surfaces upon which the various parts are painted are complex and do not correspond to those in the scene: a chair, for instance, is painted on three surfaces at right angles to one another! The painting achieves its aim (i.e., not to appear distorted) because of the constraints placed on the observer: it can only be viewed from the fixed station point with one eye, and the viewing angle is such that the extreme edges of the painted surfaces cannot be seen. Neither of these conditions apply to the pictures we normally see: we use both eyes, look at them from many different positions, and we can see the borders that separate the picture from its surroundings.

Despite the fact that trompe l'oeil pictures have been produced for hundreds of years, novel variations continue to be produced. Paintings in perspective provide an allusion to a depth not contained in their flatness – they are seen as both flat and extended in space. What would happen if a painting was produced as if it was flat but the physical surface on which it was painted was not? The artist Patrick Hughes has explored this possibility in works that are painted on protruding planes (like truncated pyramids and wedges) so that the parts that are physically close to the observer are pictorially distant. The structure is rather like Mach's card shown in Figure 7.12. That is, the lines that would converge on a flat picture plane to allude to distant objects are physically closer in these works. They appear as flat paintings until the observer moves whereupon they undergo a plastic motion that is surprising. The pictorially distant (but physically closer) parts move in the direction opposite head movement, whether this is left and right or upwards and downwards; the pictorially near (but physically more distant) parts move in the same direction as the head. These distortions occur with binocular observation and over a wide range of viewing distances.

It would seem, from this discussion, that pictures are very complex stimuli. Because we can recognise the objects represented from such a wide range of viewpoints we cannot be processing a pattern of stimulation like that from the original scene. It is more likely that we apply shape constancy to the border of the picture before processing the contents and interpreting the cues to depth. If this is the case then processes involved in picture perception must occur after those for object recognition. One consequence of this conclusion is to question whether the insights derived from picture recognition can be of any utility in furthering our understanding of object recognition.

Virtual or simulated realities

Many of the pictures we now look at are created by computers. Indeed, the technology of computer graphics is advancing so rapidly that the picture on the screen can mimic a view of a real scene and change with movements of an observer in the way a view of a real scene would. Furthermore, artificial environments can be generated by computers and presented to observers to induce the impression of exploring a three-dimensional space.

We think of simulated or virtual reality as a modern engineering enterprise, made possible by the power of computers, because they are associated with dynamic stereoscopic displays. However, the issues at the heart of simulated reality address some ancient questions of visual perception. One of these is the contrast between monocular and binocular vision, which Leonardo da Vinci repeatedly struggled with. He made many drawings of viewing with one and two eyes, and realised that the perception of depth is deficient in a painting or any other two-dimensional representation. Alberti's window provides a monocular match between a picture and a view from a single point, but what happens when two viewpoints are adopted? It is clear that Leonardo came close to representing binocular disparities, but he was unable to do so because he had no adequate theory of image formation in the eye. As was stated in Chapter 6, Wheatstone supplied this want with his invention of the stereoscope. It was the first apparatus for matching binocular incoming messages and can be considered as satisfying Leonardo's desire to imitate nature binocularly.

Neither Leonardo's painted imitation of the world from a single vantage point nor Wheatstone's paired pictures entertained the possibility of a moving perceiver. That is, the projected images themselves change as a consequence of observer movements, and the technology for capturing such changes did not then exist. It is beginning to be possible to present paired and disparate images to the two eyes and to monitor the location of the head so that the two images can be rapidly and appropriately transformed, as can sounds presented to each ear. The visual and auditory changes can be augmented by sensors which apply pressures to the surface of the skin. With such devices it is possible to present artificial or virtual worlds for observers to explore and become immersed in. Thus, the displays that are used in creating a simulated reality today slave the paired pictures to eye and head movements, so that incoming messages of a

more complicated nature can be matched to a mobile observer. Even with this addition, however, the incoming binocular messages of the simulated space cannot be matched perfectly from two flat screens, because the differential defocusing of the retinal image is not yoked to the states of convergence and accommodation of the eyes.

It is clear from the discussion above that simulated reality is by no means a novelty if we consider that pictures provide indirect access to a space other than the one occupied by the visual display. However, the increasing power of cameras and computers has added a new dimension to both simulated reality and remote vision – the visual display of current information that can be rapidly updated. Not only does this offer the opportunity to have apparently moving pictures in real time, but it also allows the viewer to interact, and control, the sequence of scenes displayed visually. The advantages of such systems are obvious for explorations of scenes that are inaccessible otherwise (like the surface of planets) or are dangerous (like the inner workings of nuclear power stations). They also permit the exploration of simulated environments.

Simulated reality is concerned with creating an imitation of the visual world. It is considered, wrongly, to be a modern engineering enterprise, made possible by the power of computers. To repeat the point made earlier, there has been very little investigation of the recognition of real, three-dimensional objects, rather than arbitrary symbolic patterns like line drawings or photographs of objects. Such patterns play an important role in human activity, in the form of alphanumeric symbols and pictures of all sorts, but are not encountered in the natural environment.

Reference notes

The representation of the space within which we live is essential to survival and it has provided one of the constant topics in philosophy, psychology and more recently computer vision. As such, almost all the books and articles we have cited in previous chapters are relevant in some respect to this chapter.

The perceptual constancies are treated in virtually every book on visual perception but it is not generally considered as a precursor of object recognition, as we argue here. In order to recognise objects it is necessary to have some indication of their orientation, size, shape and colour. Orientation is of particular significance, and it was this aspect of spatial representation that informed much of the body of Ian Howard's research. In addition to the works of his that have been cited, *Human spatial orientation* by Howard and Templeton (1966) surveyed the area and signalled the issues that warranted attention. One of the topics covered was vision with an upright image, following the experiments conducted by Stratton (1897a, b, c) and Kohler (1962). Howard and Templeton had more sympathy for the methods adopted by Stratton than the phenomenological approach of Kohler. The questions posed by optical inversion continue to be addressed by modern imaging techniques: Linden *et al.* (1999) examined four subjects wearing inverting optical devices for up to ten days. They presented psychophysical evidence to indicate

that the perceptual world remained inverted throughout the period of adaptation, and brain imaging (fMRI) indices displayed no changes in activity at early visual areas. Köhler's (1940) observation of the difficulty of recognising photographs of faces when they are upside down has spawned a vast amount of research (see Calder *et al.*, 2011). The artistic device of facial inversion not only has a much longer history – examples can be found in Roman pottery – but it has been applied with great ingenuity (see Wade, Kovács and Vidnyánszky, 2003; Wade and Nekes, 2005). Even gender and species boundaries have been crossed, as can be seen in Figure 8.2. We are most grateful to Werner Nekes for permission to reproduce these images.

Walsh and Kulikowski (1998) have edited a book entitled *Perceptual constancy* which examines many features of maintaining perception of object properties despite changing stimulation at the receptor level. The subtitle, *Why things look as they do*, relates to a question posed by the Gestalt psychologist, Kurt Koffka; Purves and Lotto (2010) have used a similar statement as the title of their book. Size constancy tends to have the lion's share of research and considerations of it stretch back to the time of Ptolemy (see Chapter 2). The relationship between the apparent size of an afterimage and the distance of the surface on which it projected is called Emmert's law, although it was described centuries earlier. Manipulations of the putative sources underlying size constancy show that when the amount of the visible surroundings and the viewing conditions are systematically reduced size constancy approaches zero. Size constancy is implicated in many illusions, like the moon illusion; Ross and Plug (2002) provide a historical account of size perception in general as well as its application to celestial illusions. The American perceptual psychologist Walt Gogel developed over many years a comprehensive theory of spatial perception, which explains constancies in terms of the interconnection between perceived size, distance and self movement (Gogel, 1990; Swanston and Gogel, 1986). Gogel and Mershon (1968) examined the distortions of apparent size induced by the Ames room. Ames devised many demonstrations that were based, essentially, on Alberti's window (see Figure 2.20). Ames' life has been chronicled by Behrens (1993). The photographs shown in Figure 8.4 were taken in Lingelbach's barn of illusions; we are grateful to Isabelle and Heinrich Bülthoff for acting as models. Details of a range of demonstrations devised by Ames are provided by Ittelson (1952). A review of shape constancy can be found in Bennett (2012) and Gilchrist (2006) provides a theory of lightness constancy.

Object recognition has accumulated a large body of research but the objects recognised are not so varied. Despite the label under which the research is conducted, the vast majority of experiments have examined representations (pictures presented on paper or screens) rather than the objects themselves (see Pinto, Cox and DiCarlo, 2008). This is perhaps because it is much easier to manipulate pictures than objects and thus to conduct experiments on such pictorial manipulations. There has been a trend in recent years to examine 'natural images', but these are often more complex pictures (of natural scenes rather than line drawings). The area in which real objects are recognised are those involving moving observers (see Findlay and Gilchrist, 2003; Land and Tatler, 2009).

Many of the characteristics of pictures and their relationships to their referents were mentioned briefly in Chapter 2, and the works cited in the Reference Notes to that chapter are relevant here. The books by Massironi (2002) and Willatts (1997) are concerned with pictures and the nature of representation, and reward reading. In addition, the books edited by Skov and Vartanian (2009) and Bacci and Melcher (2011) contain chapters that are relevant to the issues presented here; in the latter, the chapter by Melcher and Cavanagh is particularly recommended. Magritte's oeuvre has been critically assessed by Allmer (2009), and there are many other books and websites that display copies of Magritte's paintings. Canonical (or stereotypical) viewpoints refer to those that either result in a representation being more readily recognisable or preferential; they have been studied extensively by Stephen Palmer and are discussed in the context of art and aesthetics by Palmer, Schloss and Sammartino (2012). Some of the perceptual aspects of cave painting are discussed in Melcher and Wade (2006; Wade and Melcher, 2006).

Brunswik (1956) reported a series of studies on the perception of schematic faces. The same oval outline contained four simple structures – single lines for the nose and mouth, together with elliptical eyes. The separation of the eyes, the length of the nose, and the height of the mouth were varied systematically to yield 189 faces. Observers were requested to rank the faces according to characteristics like narrow-minded, joyful, embittered, sad, and sarcastic, and there was reasonable agreement over which faces had these attributes. Schematic drawings like these are now rarely used in experiments on face perception due to the advances in imaging methods (see Calder *et al.*, 2011).

Artists have revelled in producing ambiguous and distorted pictures and this has also provided fascination for visual scientists. The Necker cube is one of the most frequently reproduced examples of an ambiguous figure, although it was manipulated artistically long before it was examined scientifically. Details of Necker's life can be found in Wade *et al.* (2010). Designs similar to those in Figure 8.14 have been found in the House of Dionysius, Cyprus (Kondoleon, 1995), and in the mosaics of Pompeii (Lingelbach, 2008). These systematic manipulations of simple geometrical shapes to represent solid forms are testament to the knowledge of perceptual organisation in the designers of the mosaics (see Wade, 2012b). Illustrations of trompe l'oeil art are presented in Giusti (2009).

9 Summary and conclusions

The study of perception has a long history, as long as the history of science itself. The framework for science and for the study of perception was outlined by Greek philosophers, but we have concentrated upon the advances made in the modern era, following the scientific revolution in the seventeenth century. Our current theories of perception have been shaped by ideas in physics, art, physiology, biology, philosophy, psychology, and more recently computer science. Physics contributed an understanding of the nature of light and of the laws of optical transmission in air and in the eye. The former led to the recognition that vision is not due to the emission of light from the eye, and the latter made it possible to understand how objects in the external world were represented as images on the retinal surface. The laws of optics were also employed in artistic representation: the formation of an image on a transparent screen at a fixed distance in front of the eye will produce a picture in linear perspective; if the screen is considered to lie at the back of the eye then a similar pattern will be produced (though inverted and reversed with respect to the external picture). When this similarity between image formation in the eye, in cameras, and in art was appreciated the problem of perception was thought of in terms of extracting information from a two-dimensional image. The retinal image was conceived in terms of a static picture, which has had profound effects upon both theory and experiment. It has meant that pictorial representations of objects could be considered as adequate experimental substitutes for the objects themselves, and that the dynamic aspects of vision could be too readily ignored. Because our eyes are in constant motion, the retinal image is an abstraction, not actually present in the eye unless a scene is viewed for a very brief interval, too short for the eyes to move. Anatomists examined the structure of the retina and of the optic nerves. Physiologists conjectured how the brain could distinguish between neural signals arising from different sense organs.

Towards the end of the nineteenth century, specific cortical regions were found which received neural signals from the different sense organs. The biological perspective, following Darwin's theory, placed humans in a closer relationship to other animals and so added to the relevance of experiments on animal perception to interpretations of human perception. It emphasised developmental processes in both behaviour and perception; and it also examined instinctive and learned behaviours in more detail. Issues of innate and learned processes in perception

have provided constant conflict within philosophy. On the one hand, the nativists argued that certain aspects of perception (like space and time) were inborn, on the other the empiricists proposed that we derive our knowledge of the world through the senses and we learn to perceive spatial attributes by a process of association. These distinctions remain with us and have permeated psychological theories generally. Psychology has provided the study of perception with its basic experimental methodologies. It was the harnessing of psychophysical methods and the invention of a novel one for studying consciousness (analytic introspection) that led to the formation of psychology as an independent discipline. The psychophysical methods have stood the test of time (with constant improvements), but analytic introspection has fared less well: its demise saw the emergence of Gestalt and Behaviourist psychologies in the early part of the twentieth century, with the former showing more concern with perceptual matters than the latter.

Gestalt theory drew upon facts of magnetic and electrical fields in physics and speculated about similar field influences in the brain. That is, it considered that both perception and its underlying physiology were modelled on organising principles that could not be reduced to the operations of simplified elements. The Gestalt psychologists emphasised the holistic aspects of perception rather than the analytic: they did not believe that perception could be explained by reducing either stimuli or responses to their simplest levels. This contrast between holistic and analytic approaches to perception is one that remains in theories of perception. Ironically, the subsequent advances in the neurophysiology of vision have been based on precisely the reductionist ideas that the Gestalt psychologists rejected. The detailed study of the structure and function of nerve cells has suggested the ways in which complex patterns of light striking the retina are analysed in terms of simplified features that they contain. Cortical cells could be excited by chromatic or spatiotemporal changes in the pattern of light falling over particular regions of the retina. Thus, features of the stimulus (like retinal location, contour orientation, direction of motion, and colour) are thought to be extracted by single cells in the visual cortex of higher mammals, and they are preferentially processed at subsequent cortical sites. The presence of such feature detectors and cortical maps has been related to certain simple phenomena like visual aftereffects and illusions. However, the range of phenomena that has been successfully interpreted in neurophysiological terms remains limited, despite undoubted technological advances such as magnetic resonance imagery (MRI) and other techniques for locating activity in the brain. These methods provide indirect evidence of neural activity (e.g. by increased oxygen uptake) and are able to localise brain responses to different categories of stimuli, at the cost of losing information about the timing of such responses. Attempts to identify the brain region responsible for a psychological process can be contrasted with an emphasis on the overall activity of the brain, as is the case with the study of neural nets. Computer models of the way in which diffuse interconnected networks can be shaped by patterns of input and feedback from outcomes may provide a basis for returning to a more holistic approach to

perception. Such networks have capacities for learning which are not localised in particular units, but which are a property of the net as a whole.

We know much more about the visual neurophysiology of cats and monkeys when presented with simple patterns than we do of humans. On the other hand, we have a great body of knowledge about human perception under similar simplified circumstances. In recent years this has been extended by the use of patterns of parallel lines (gratings). It seems as though such patterns are processed at early stages in vision, and there is evidence that spatial frequency is another stimulus dimension that is extracted by some single cortical cells. This has been a case of psychophysical experiments with human subjects leading to the neurophysiological search for a specific class of detectors in monkey visual cortex.

We still do not know how these feature detectors are involved in perception generally. This is largely because in almost all the neurophysiological experiments that are conducted with experimental animals the eye muscles are paralysed to avoid any eye movements during the presentation of stimuli. In our terms these studies are concerned with stimulus processing at the retinocentric level. That is, the representation of features like contour orientation, spatial frequency, direction of motion and colour is with respect to the retinal coordinate system alone. Such a representation is unlikely to be adequate for vision in an animal with moving eyes, because retinocentric displacements could be a consequence of object movement or eye movement.

Perception of objects in the world involves seeing where they are, when they occur and what they are. That is, we need to determine their location, whether they are moving, and to recognise them. These characteristics are defined relative to a frame of reference, so that statements about an object's motion, for example, are typically made with respect to the surface of the earth. Perception of object characteristics needs to conform to the frame of reference within which behaviour takes place, which will also normally be the surface of the earth. Such perceptions and associated actions can be termed geocentric. An active perceiver, whether human or animal, must recover the geocentric properties of objects, despite self-motion. We have described the frames of reference relative to which information can be represented, and from which geocentric information can be derived. The first, retinocentric, level is the one that is studied by neurophysiologists: information is coded in terms of the coordinate system of the retina. Retinocentric information describes image characteristics with respect to retinal coordinates, but it is not necessarily localised in the retina. The fact that a response occurs in the visual cortex, or elsewhere, need not preclude it from being retinocentric in character. The term retinocentric expresses the nature of the information, and not its anatomical site.

This level is essential for vision but it cannot serve as a basis for object perception because we have two eyes and both of them move. The next frame of reference uses an integrated binocular signal from the two eyes together with their movements to provide an egocentric frame of reference, the origin of which lies between the eyes. We perceive the directions of objects with respect to the egocentre. Visual direction alone would not enable us to locate objects

in space; in order to do this we also need to determine the distance objects are away from us. There are many sources of information for the relative distances between objects – which is nearer or further – but these need to be anchored by information for object distance before objects can be adequately located in space. If information for the observer's own movements is incorporated, then objects will be represented in the geocentric frame of reference. When we have derived a geocentric representation of objects, we are in a position to behave appropriately with respect to them. The operation of these frames of reference can be more fully illustrated in the context of motion perception.

Perceiving whether, and to what extent, an object is moving depends on many factors. Consideration of retinal stimulation alone cannot account for geocentric visual perception. Many other sensory systems are involved in the recovery of the location and movement of objects in space. For example, information from the vestibular system is used to determine accelerations of the head in three dimensions, and this is essential if the pattern of retinal change arising from a head movement is to be interpreted unambiguously. The integration of these sensory systems constitutes visual perception, and it would be wrong to suppose that visual perception can be understood solely in terms of light reaching the eye. Initial registration of retinocentric motion in each eye gives rise to a single binocular (cyclopean) retinocentric signal. This expresses motion with respect to a point located between the eyes. Since the two eyes move by equal amounts when version occurs, a single signal for eye movement can be combined with the binocular retinocentric signal to give an egocentric representation. This carries information for changes in angular direction with respect to the observer, despite movements of the eyes. It will correctly express changes in object direction, but will only represent the position of objects in the environment if the observer is stationary. If the observer's head moves then the eyes necessarily move, too. This causes displacement of the retinal image, in a way which cannot be interpreted unless the extent of self-motion and the distance of the object are known. Scaling the egocentric information by perceived distance gives the three-dimensional location of objects with respect to the observer. This can be corrected for the effects of self-movement, since both are expressed in terms of three-dimensional displacements, to produce a geocentric representation of the movement of objects in the environment.

We can generate a geocentric representation of space and motion under the conditions described above, although they are rarely encountered in the natural environment. Objects are seldom seen in isolation; other background objects will be in view, and they might undergo different transformations relative to the observer. Moving objects are seen in the context of others that are typically supported on the ground plane and remain in a fixed position with respect to it. Thus, there is another frame of reference that is of vital significance to perception and visual-motor control generally. It concerns the location and motion of one part of a visual scene with respect to others. Since it refers to visual pattern interactions we have called it a patterncentric frame of reference. Thus, if one object moves with respect to others (that remain in the same spatial relations with respect

to one another) then the object is seen to be moving. This is the situation that obtains almost all the time in our interaction with the world. For example, cars are seen as moving because the features surrounding them, like trees and buildings, remain in the same relation to one another. The trees and buildings provide a patterncentric frame of reference relative to which the displacements of the car are assigned. Stars, on the other hand, are not so firmly anchored in perceptual space: they can appear to be stationary or to move, depending upon their visible background. They appear stationary on a cloudless night, but can seem to move when clouds pass by them. This indicates the manner in which the patterncentric process operates: when the stars are seen behind moving clouds they appear to move in the opposite direction to the clouds. This is called induced motion, and it reflects the influence of a larger visual pattern on the visibility of parts enclosed within it. The clouds provide a patterncentric frame of reference, and are treated as stationary, whereas the displacement of the stars with respect to them is treated as motion. Under almost all normal conditions the patterncentric frame of reference will correspond to that operating geocentrically. However, they can be separated in the laboratory by removing or manipulating the visible background. The examples of induced motion given above and in Chapter 7 involved isolating and moving visible backgrounds. Our sensitivity to relational motion is far more acute than that for uniform motion, as was indicated in Chapter 7. Induced motion is also the phenomenon exploited in many video games, where a central pattern (like a car) appears to move in the opposite direction to the moving background. Such induced motion is rarely complete because the monitor on which the motion is presented is both visible and in the same position relative to other objects (like the walls of the room).

At this level of analysis, the nature of the information required for a geocentric representation is defined, rather than its sources. The concern is with the rules and relationships that permit a geocentric representation of the external world by an observer who can move in the environment, and whose sense organs can move with respect to the body. A geocentric representation is required for effective action in the environment, since it conveys information which is independent of the observer's movements. A geocentric representation has the properties required for perceptual constancies, such as orientation, size, shape and location. It will not change despite changes in viewing distance or angle of regard. It will therefore provide a basis for object recognition, since objects are represented in a manner which is consistent even when the conditions of observation change. If a geocentric representation is stored in memory, it can be matched to the same object when encountered on a subsequent occasion, because it is independent of the observer and the circumstances of observation.

There may, of course, be errors in the process of recovering a geocentric description. Relational motions, eye movements, self-movements or distance may not be accurately detected, and image displacements on the retina may be more or less than the signal to which they give rise. If so, then the perception of orientation, shape, size or movement may no longer accord with their physical dimensions, and object recognition may fail.

Geocentric representations may be employed to carry out tasks that are expressed in terms of other coordinate systems. For example, someone may be asked to state the direction of an object with respect to their head, or some other part of their body. Alternatively, it may be necessary to judge the location of one object with respect to another. These tasks can be accomplished on the basis of geocentric information. Although the required frame of reference is no longer explicit, it can be recovered when needed if information about the articulation of the eyes, head and body is preserved. It is important to distinguish the frame of reference used to describe an activity from that which determines an observer's spatial perception.

All perception is necessarily geocentric, but it may not seem so under restricted laboratory conditions. Stimuli may be exposed for intervals too short for the eyes to move during their presentation. If very briefly exposed stimuli are used, eye movements will not take place, and perception will correspond to the egocentric level. It might be thought that in this case perception would be retinocentric, but binocular combination is unavoidable, even when one eye is covered or non-functional. Egocentric directions and movements are always based upon binocular signals, at least in humans and other species with overlapping monocular visual fields. If the observer remains stationary, then the importance of perceived distance and perceived self-motion in spatial vision will not be appreciated. A high proportion of studies of human spatial perception incorporate restrictions that eliminate one or more of these factors.

Although this discussion has been expressed in terms of the fundamental requirements for human perception, it is interesting to consider the similarities, and differences, that may apply to other species. Few animals have large overlapping binocular fields of view, and some have more than two eyes. In some crustaceans and insects, eyes are mounted on the end of flexible stalks which can move in three dimensions relatively independently of the rest of the body. Despite these complications, it is still necessary to obtain a geocentric representation of the world. Even a cursory examination of animal behaviour shows that most actions require information about the location in space of salient objects like predators, prey, mates or offspring. Although we have considerable knowledge about the physiology of vision in many species, we know much less about the logic of the perceptual processes which they serve. One point at any rate is clear: geocentric perception is not necessarily based on knowledge of the world expressed in linguistic symbols. If this were so, only humans and perhaps the higher primates would be able to act effectively in the natural environment. Of course, a geocentric representation may be achieved in quite different ways by different species, just as a given arithmetical problem may have a number of equally valid routes to the correct answer. Nevertheless, the underlying requirements are the same; to obtain from the senses information which is capable of guiding the activities required for survival.

The role of cognitive processes in perception has been exceptionally difficult to establish. To an extent, this has become a matter of theoretical dispute. Those who have followed Gibson's approach have regarded perception as wholly

independent of cognition, and determined entirely by a direct response to patterns of stimulation without any intervening process. By contrast, others have claimed that perception is akin to high-level cognitive processes like problem solving and hypothesis formation. What we see would then be determined at least as much by what we know as by what is there to be seen. It is in the context of arbitrary stimuli, like pictures and written script, that the cognitive approaches to perception are at their strongest. We learn to attach significance to certain shapes that formerly were of no importance to us. However such learning processes are not confined to humans. Pigeons can discriminate between letters of the alphabet, and many artificial seeing machines can carry out industrial inspection tasks. Our metaphorical guide dog is also able to learn the significance of initially arbitrary objects and act on them, like stopping in front of overhanging barriers that it, but not a human, could walk under.

The goal of this book has been to concentrate on the functions which a perceptual system must perform, and to show how the understanding of these has developed in the modern era. It is very clear that many of the theoretical issues defined in the eighteenth, nineteenth and twentieth centuries are still not resolved, and they persist despite great advances in techniques for manipulating visual stimulation. Nevertheless, vision is an endlessly fascinating object of study, drawing as it does on such diverse disciplines for insight and advance. Even if the final goal is not in sight, the journey is worth the effort.

References

Addams, R. (1834) An account of a peculiar optical phænomenon seen after having looked at a moving body. *London and Edinburgh Philosophical Magazine and Journal of Science*, 5, 373–374.

Aguilonius, F. (1613) *Opticorum libri sex. Philosophis juxta ac mathematicis utiles*. Antwerp: Moretus.

Alais, D., and Blake, R. (eds.) (2005) *Binocular rivalry and perceptual ambiguity*. Cambridge, MA: MIT Press.

Aleman, A. and Larøi, F. (2008) *Hallucinations: The science of idiosyncratic perception*. Washington DC: APA Books.

Allmer, P. (2009) *René Magritte – beyond painting*. Manchester: University of Manchester Press.

Ash, M. G. (1995) *Gestalt psychology in German culture, 1890–1967: Holism and the quest for objectivity.* Cambridge, MA: Cambridge University Press.

Atkinson, J. (2002) *The developing visual brain*. Oxford: Oxford University Press.

Aubert, H. (1865) *Physiologie der Netzhaut*. Breslau: Morgenstern.

Aubert, H. (1887): Die Bewegungsempfindungen. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, 40, 459–480.

Bacci, F., and Melcher, D. (eds.) (2011) *Art and the senses*. Oxford: Oxford University Press.

Bartisch, G. (1583) *ΟΦΘΑΛΜΟΥΛΕΙΑ. Das ist, Augendienst*. Dresden: Stockel.

Bartlett, F. C. (1932) *Remembering: A study in experimental and social psychology*. Cambridge, UK: University Press.

Bear, M. F., Connors, B. and Paradiso, M. (2006) *Neuroscience: Exploring the brain.* 3rd edn. London: Lippincott, Williams and Wilkins.

Behrens, R. R. (1993) *The man who made distorted rooms: A chronology of the life of Adelbert Ames Jr.* University of North Iowa: Published by the author.

Bennett, D. J. (2012) Seeing shape: Shape appearances and shape constancy. *British Journal of the Philosophy of Science*, 63, 487–518..

Berkeley, G. (1709) *An essay towards a new theory of vision*. Dublin: Pepyat.

Berrios, G. E. (2005) On the fantastic apparitions of vision by Johannes Müller. *History of Psychiatry*, 16, 229–246.

Besharse, J. C., and Bok, D. (2011) *The retina and its disorders*. Oxford: Elsevier.

Billock, V. A., and Tsou, B. H. (2011) To honor Fechner and obey Stevens: Relationships between psychophysical and neural nonlinearities. *Psychological Bulletin*, 137, 1–18.

Birch, T. (ed.) (1966) *Robert Boyle. The works*. Hildesheim: Olms.

Blakemore, C. (1973) The baffled brain. In R. L. Gregory and E. H. Gombrich (eds.) *Illusion in nature and art*. London: Duckworth. pp. 9–47.

Blundell, B. (2011) *3D displays and spatial interaction: exploring the science, art, evolution and use of 3D technologies*. Auckland, NZ: Walker & Wood.

Bonnet, C. (1760) *Essai analytique sur la facultés de l'âme*. Copenhague: Philibert.

Boring, E. G. (1942) *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century.

Braun, M. (1992) *Picturing time: The work of Étienne-Jules Marey (1830–1904)*. Chicago: University of Chicago Press.

Brewster, D. (1832) *Letters on natural magic addressed to Sir Walter Scott, Bart*. London: Murray.

Brewster, D. (1855) *Memoires of the life, writings, and discoveries of Sir Isaac Newton*. Edinburgh: Constable.

Brewster, D. (1856) *The stereoscope. Its history, theory, and construction*. London: Murray.

Bridge, H. (2011) Mapping the visual brain: How and why. *Eye*, 25, 291–296.

Bridgeman, B. (2010) How the brain makes the world appear stable. *i-Perception*, 1, 69–72. http://i-perception.perceptionweb.com/fulltext/i01/i0387.pdf

Broadbent, D. E. (1958) *Perception and communication*. London: Pergamon Press.

Brunswik, E. (1934) *Wahrnehmung und Gegenstandwelt. Grundlegung einer Psychologie vom Gegenstand her*. Leipzig: Deuticke.

Brunswik, E. (1956) *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Burgess, N. (2006) Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences*, 10, 551–557.

Burton, H. E. (1945) The optics of Euclid. *Journal of the Optical Society of America*, 35, 357–372.

Cahan, D. (ed.). (1993) *Hermann von Helmholtz and the foundations of nineteenth-century science.* Berkeley, CA: University of California Press.

Calder, A. J., Rhodes, G., Johnson, M. H., and Haxby, J. V. (eds). (2011) *The Oxford handbook of face perception.* Oxford: Oxford University Press.

Chevreul, M. E. (1839) *De la loi du contraste simultané des couleurs*. Paris: Pitois-Levrault.

Clegg, B. (2001) *Light years. An exploration of mankind's enduring fascination with light*. London: Piatkus.

Clément, G. (2011) *Fundamentals of space medicine*. 2nd edn. New York: Springer.

Colenbrander, A. (2001) Measuring vision and vision loss. In *Duane's Clinical Ophthalmology* Vol. 5 Eds. W. Tasman, E. A. Jaeger. Philadelphia: Lippincott, Williams & Wilkins.

Copernicus, N. (1978) *On the revolutions*. Trans. and ed. J. Dobrzycki. London: Macmillan.

Coren, S., Ward, L. M., and Enns, J. T. (2004) *Sensation and perception*. 6th edn. New York: Wiley.

Cowles, M. P. (1989) *Statistics in psychology. An historical perspective*. Hillsdale, NJ: Erlbaum.

Craik, K. J. W. (1943) *The nature of explanation*. Cambridge, UK: University Press.

Craik, K. J. W. (1966) *The nature of psychology. A selection of papers, essays and other writings by the late Kenneth J. W. Craik*. Cambridge, UK: University Press.

Crombie, A. C. (1964) Kepler: De Modo Visionis. In *Mélange Alexandre Koyré I. L'Aventure de la Science*. Paris: Hermann, pp. 135–172.

Crone, R. A. (1992) *Licht-kleur-ruimte*. Dordrecht: Kluwer.

Cutting, J. E. (1986) *Perception with an eye for motion*. Cambridge, MA: MIT Press.

Daini, R., Wenderoth, P., and Smith, S. (2003) Visual orientation illusions: Global mechanisms involved in hierarchical effects and frames of reference. *Perception & Psychophysics*, 65, 770–778.

Dalton, J. (1798) Extraordinary facts relating to the vision of colours: with observations. *Memoirs of the Literary and Philosophical Society of Manchester*, 5, 28–45.

Danziger, K. (1990) *Constructing the subject: Historical origins of psychological research*. Cambridge: Cambridge University Press.

Darrigol, O. (2012) *A history of optics from Greek antiquity to the nineteenth century*. Oxford: Oxford University Press.

Darwin, E. (1794) *Zoonomia; Or, the laws of organic life*. Vol. 1. London: Johnson.

Daw, N. (2012) *How vision works: The physiological mechanisms behind what we see.* Oxford: Oxford University Press.

Daw, N. W. (2006) *Visual development*. (2nd edn.) New York: Springer.

Daxecker, F. (2004) *The physicist and astronomer Christoph Scheiner: Biography, letters, works*. Innsbruck: Leopold-Franzens-University of Innsbruck.

Day, R. H., and Wade, N. J. (1969) Mechanisms involved in visual orientation constancy. *Psychological Bulletin*, 71, 33–42.

Descartes, R. (1637/1902). *La dioptrique*. In C. Adam and P. Tannery (eds.) *Oeuvres de Descartes*, Vol. 6. (pp. 81–228). Paris: Cerf.

Descartes, R. (1662) *De homine figuris et latinitate*. Trans. F. Schuyl. Lugduni Batavorum: Leffen & Moyardum.

Descartes, R. (1965/1637) *Discourse on method, optics, geometry, and meteorology*. Trans. P. J. Olscamp. Indianapolis: Bobbs-Merrill.

Donahue, W. H. (2000) *Johannes Kepler. Optics: Paralipomena to Witelo & optical part of astronomy.* Santa Fe, NM: Green Lion Press.

Donders, F. C. (1864) *On the anomalies of accommodation and refraction of the eye*. London: New Sydenham Society.

Ellis, W. D. (ed.) (1938) *A source book of Gestalt psychology*. London: Routledge and Kegan Paul.

Epstein, W., and Rogers, S. (eds.) (1995) *Perception of space and motion*. New York: Academic Press.

Fahle, M., and Greenlee, M. (2003) *The neuropsychology of vision*. Oxford: Oxford University Press.

Faraday, M. (1831) On a peculiar class of optical deception. *Journal of the Royal Institution*, 1, 205–223.

Farah, M. J. (2004) *Visual agnosia.* (2nd edn.) Cambridge, MA: MIT Press.

Fechner, G. T. (1860) *Elemente der Psychophysik*. Leipzig: Breitkopf and Härtel.

Fechner, G. T. (1966) *Elements of psychophysics*. Vol. 1. Trans. H. E. Adler, Ed. D. H. Howes and E. G. Boring. New York: Holt, Rinehart and Winston.

Findlay, J. M. and Gilchrist, I. D. (2003) *Active vision. The psychology of looking and seeing*. Oxford: Oxford University Press.

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., and Macleod, D. I. (2003) Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, 6, 909–910.

Finger, S. (1994) *Origins of neuroscience. A history of explorations into brain function*. New York: Oxford University Press.

Ford, B. J. (2007) Enlightening neuroscience: microscopes and microscopy in the eighteenth century. In H. Whitaker, C. U. M. Smith, and S. Finger (eds.) *Brain, mind, and medicine: Essays in 18th century neuroscience*. New York: Springer. pp. 29–41.

Frackowiak, R. S. J, Friston, K. J., Frith, C. D., Dolan, R. J., and Mazziotta, J. C. (2004) *Human brain function*. 2nd edn. London: Academic Press.

Freeman, T. C. A., Champion, R. A., and Warren, P. A. (2010) A Bayesian model of perceived head-centred velocity during smooth pursuit eye movement. *Current Biology*, 20, 757–762.

Frisby, J. P and Stone, J. V. (2010) *Seeing. The computational approach to biological vision*. 2nd edn. London: MIT Press.

Gardner, H. (1987) *The mind's new science. A history of the cognitive revolution*. New York: Basic Books.

Geier, J., and Hudák, M. (2011) Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold – a psychophysical refutation. *PLoS ONE*, 6(10): e26062. doi:10:1371/journal.pone.0026062.

Gibson, J. J. (1950) *Perception of the visual world*, Boston: Houghton Mifflin.

Gibson, J. J. (1966) *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1968) What gives rise to the perception of motion? *Psychological Review*, 75, 335–346.

Gibson, J. J. (1979) *The ecological approach to visual perception*, Boston: Houghton Mifflin.

Gilchrist, A. (2006) *Seeing black & white*. Oxford: Oxford University Press.

Giusti, A. (2009) *Art and illusions: Masterpieces of trompe l'oeil from antiquity to the present day*. Florence: Mandragora.

Goethe, J. W. von (1810/1970). *Theory of colours*. Cambridge, MA: MIT Press.

Gogel, W. C. (1990) A theory of phenomenal geometry and its applications. *Perception & Psychophysics*, 48, 105–123.

Gogel, W. C. and Mershon, D. H. (1968) The perception of size in a distorted room. *Perception & Psychophysics*, 4, 26–28.

Goldstein, E. B. (2009a). *Sensation and perception*. 8th edn. Andover, UK: Cengage.

Goldstein, E. B. (2009b). *Encyclopedia of perception*. Thousand Oaks, CA: Sage.

Gombrich, E. H. (1960) *Art and illusion. A study in the psychology of pictorial representation*. London: Phaidon.

Goodale, M. A. and Milner, A. D. (1992) Separate pathways for perception and action. *Trends in Neuroscience*, 15, 20–25.

Goodale, M. A. and Milner, A. D. (2005) *Sight unseen. An exploration of conscious and unconscious vision*. Oxford: Oxford University Press.

Gordon, I. E. (2004) *Theories of visual perception*. 3rd edn. Hove, UK: Psychology Press.

Gregory, R. L. (1997) *Eye and brain: The psychology of seeing*. 5th edn. Oxford: Oxford University Press.

Gregory, R. L. (2006) Bayes Window (4): Table of illusions. *Perception*, 35, 431–432.

Gregory, R. L. (2009) *Seeing through illusions*. Oxford: Oxford University Press.

Grimaldi, F. M. (1665) *Physico-mathemis de lumine coloribis et iride*. Bononiae.

Gross, C. G. (1998) *Brain, vision, memory: Tales in the history of neuroscience*. Cambridge, MA: MIT Press.

Guigon, E. and Pierre, A. (2005) *L'œil moteur. Art optique et cinétique, 1950–1975*. Musées de Strasbourg: Strasbourg.

Hall, T. S. (1972) *Treatise of man. René Descartes*. Cambridge, MA: Harvard University Press.

Hammond, J. H. (1981) *The camera obscura. A chronicle*. Bristol: Hilger.

Harris, J. (1775) *A treatise of optics: Containing elements of the science; in two books*. London: White.

Harvard Medical School (2001) *The aging eye*. New York: Simon & Schuster.
Hebb, D. O. (1949) *Organization of behavior: A neuropsychological theory*. New York: Wiley.
Helmholtz, H. (1867) *Handbuch der physiologischen Optik*. In G. Karsten, Ed. *Allgemeine Encyklopädie der Physik*. Vol. 9. Leipzig: Voss.
Helmholtz, H. (1873) *Popular lectures on scientific subjects*. First series. Trans. E. Atkinson. London: Longmans, Green.
Helmholtz, H. (1881) *Popular lectures on scientific subjects*. Second series. Trans. E. Atkinson. London: Longmans, Green.
Helmholtz, H. (1925/2000). *Helmholtz's Treatise on physiological optics*. 3 Vols. Trans. J. P. C. Southall. Bristol: Thoemmes.
Herbert, S. (ed.) (2000) *A history of pre-cinema.* Vol. 1. London: Routledge.
Hering, E. (1868/1942) *Spatial sense and movement of the eye,* trans. A. Radde, Baltimore: American Academy of Science.
Holland, H. C. (1965) *The spiral after-effect*. Oxford: Pergamon.
Hooke, R. (1665) *Micrographia: or some physiological descriptions of minute bodies made by magnifying glasses with observations and inquiries thereupon*. London: Martyn and Allestry.
Hopkins, A. A. (1898) *Magic: Stage illusions and scientific diversions, including trick photography*. New York: Munn.
Howard, I. P. (1982) *Human visual orientation*. Chichester, UK: Wiley.
Howard, I. P. (1993) Spatial vision within egocentric and exocentric frames of reference. In S. R. Ellis, M. Kaiser and A. J. Grunwald (eds.) *Pictorial communication in virtual and real environments*. 2nd edn. London: Taylor and Francis. pp. 338–358.
Howard, I. P. (1999) The Helmholtz–Hering debate in retrospect. *Perception*, 28, 543–549.
Howard, I. P. (2012) *Perceiving in depth. Vol. 1. Basic mechanisms*. New York: Oxford University Press.
Howard, I. P., and Rogers, B. (2012) *Perceiving in depth. Vol. 2. Stereoscopic vision*. New York: Oxford University Press.
Howard, I. P. and Templeton, W. B. (1966) *Human spatial orientation*. London: Wiley.
Howard, I. P., Hu, G., Saxe, R., and James, E. Z. (2005) Visual orientation in a mirror world tilted 90 degrees. *Perception*, 34, 7–15.
Huang, D., Kaiser, P. K., Lowder, C. Y., and Traboulsi, E. I. (2005) (eds.) *Retinal imaging*. Philadelphia: Mosby Elsevier.
Hubel, D. H. and Wiesel, T. N. (2005) *Brain and visual perception. The story of a 25-year collaboration*. Oxford: Oxford University Press.
Humboldt, A. von (1850) *Kosmos. Entwurf einer physischen Weltbeschreibung, Vol. 3*, Stuttgart: Cotta.
Humphreys, G. W., and Riddoch, M. J. (1986) *To see but not to see. Case study of visual agnosia*. Hove, UK: Psychology Press.
Hunt, D. M., Dulai, K. S., Bowmaker, J. K. and Mollon, J. D. (1995) The chemistry of John Dalton's color blindness. *Science*, 267, 984–988.
Huygens, C. (1690) *Traité de la lumiere*. Leiden: Vander.
Huygens, C. (1912) *Treatise on light*. Trans. S. P. Thompson. London: Macmillan.
Ilardi, V. (2007) *Renaissance vision from spectacles to telescopes*. Philadelphia: American Philosophical Society.
Ings, S. (2007) *The eye. A natural history*. London Bloomsbury.
Ittelson, W. H. (1952) *The Ames demonstrations in perception*. Princeton, NJ: Princeton University Press.

Jacobs, G. H. (2009) Evolution of colour vision in mammals. *Philosophical Transactions of the Royal Society of London B*. 364, 2957–2967.

James, W. (1890) *Principles of psychology. Vol. 1*. New York: Holt.

Johansson, G. (1973) Visual perception of biological motion and a model for its analysis, *Perception and Psychophysics*, 14, 201–211.

Johansson, G. (1975) Visual motion perception, *Scientific American*, 232, June, 76–89.

Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (2000) *Principles of neural science*. 4th edn. New York: McGraw-Hill.

Kanwisher, N., and Barton, J. J. S. (2011) The functional architecture of the face system: Integrating evidence from fMRI and patient studies. In A. J. Calder, G. Rhodes, M. H. Johnson, and J.V. Haxbey (eds) *Oxford handbook of face perception*. Oxford: Oxford University Press.

Kanwisher, N., and Duncan, J. (eds.) (2004) *Functional neuroimaging of visual cognition (Attention and Performance series XX)*. New York: Oxford University Press.

Kemp, M. (1990) *The science of art. Optical themes in Western Art from Brunelleschi to Seurat*. Hartford, CT: Yale University Press.

Kemp, M. (2001) *Leonardo on painting*. Newhaven, CT: Yale University Press.

Kepler, J. (1604) *Ad Vitellionem paralipomena*. Frankfurt: Marinium and Aubrii.

Kepler, J. (1611) *Dioptrice*. Augsburg: Franci.

Kirby, J. (1755) *Dr Brook Taylor's Method of Perspective made easy, both in Theory and Practice. In Two Books*, 2nd edn. Ipswich: published by the author.

Kohler, I. (1962) Experiments with goggles. *Scientific American*, 206 (5), 62–86.

Köhler, W. (1930) *Gestalt psychology*. London: Bell.

Köhler, W. (1940) *Dynamics in psychology*. London: Faber and Faber.

Kondoleon, C. (1995) *Domestic and divine: Roman mosaics in the House of Dionysos.* Ithaca, NY: Cornell University Press.

Land, M. F. and Nilsson, D.-E. (2001) *Animal eyes*. Oxford: Oxford University Press.

Land, M. F. and Tatler, B. W. (2009) *Looking and acting: Vision and eye movements during natural behaviour.* Oxford: Oxford University Press.

Leonardo da Vinci (1721) *A treatise of painting*. London: Senex and Taylor.

Lindberg, D. C. (1976) *Theories of vision from Al-Kindi to Kepler*. Chicago: University of Chicago Press.

Linden, D. E. J., Kallenbach, U., Heinecke, A., Singer, W., and Goebel, R. (1999) The myth of upright vision. A psychophysical and functional imaging study of adaptation to inverting spectacles. *Perception,* 28, 469–481.

Lingelbach, B. (2008) Der Necker-Würfel – oder was ist vorne. *Klinische Monatsblätter für Augenheilkunde*, 225, 253–254.

Liversedge, S., Gilchrist, I., and Everling, S. (eds.) (2011) *The Oxford handbook of eye movements*. Oxford: Oxford University Press.

Livingstone, M. (2002) *Vision and art. The biology of seeing*. New York: Abrams.

Locke, J. (1690) *An essay concerning human understanding*. Oxford: Clarendon.

Locke, J. (1694) *An essay concerning human understanding*. 2nd edn. Oxford: Clarendon.

Lucretius (1975) *De Rerum Natura*. Trans. W. H. D. Rouse. Cambridge, MA: Harvard University Press.

Mach, E. (1866) Ueber die physiologische Wirkung raümlich vertheilter Lichtreize. (Dritte Abhandlung) *Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, 54, 393–408.

Mannoni, L. (2000) *The great art of light and shadow. Archeology of the cinema*. Trans. R. Crangle. Exeter: University of Exeter Press.

Mannoni, L., Nekes, W., and Warner, M. (eds.) (2004) *Eyes, lies and illusions*. London: Hayward Gallery.

Marr, D. (1982) *Vision. A computational investigation into the human representation and processing of visual information*. New York: Freeman.

Marr, D. (2010) *Vision. A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press.

Massironi, M. (2002) *The psychology of graphic images. Seeing, drawing, communicating*. Trans. N. Bruno. Mahwah, NJ: Erlbaum.

Mather, G. (2008) *Foundations of sensation and perception*. 2nd edn. Hove, UK: Psychology Press.

Mather, G., and Murdoch, L. (1994) Gender discrimination in biological motion displays based on dynamic cues, *Proceedings of the Royal Society of London*, B, 258, 273–279.

Mather, G., Verstraten, F. and Anstis, S. (eds.) (1998) *The motion after-effect: a modern perspective*. Cambridge, MA: MIT Press.

Matthews, N., Meng, X., Xu, P., and Qian, N. (2003) A physiological theory of depth perception from vertical disparity. *Vision Research*, 43, 85–99.

May, M. T. (1968) *Galen. On the usefulness of the parts of the body*. Ithaca, NY: Cornell University Press.

McCoun, J., and Reeves, L. (2010) *Binocular vision: development, depth perception and disorders*. Hauppauge, NY: Nova Science.

Melcher, D., and Cavanagh, P. (2011) Pictorial cues in art and visual perception. In F. Bacci and D. Melcher (eds.) (2011) *Art and the senses*. Oxford: Oxford University Press. pp. 359–394.

Melcher, D., and Wade, N. J. (2006) Cave art interpretation. II. *Perception*, 35, 719–722.

Miller, G. A. (2003) The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, 7, 141–144.

Miller, G., Gallanter, E, and Pribram, K. (1960) *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston.

Miller, P. E. and Murphy, C. J. (1995) Vision in dogs. *Journal of the Veterinary Medical Association*, 207, 1623–1634.

Milner, A. D., and Goodale, M. A. (1995) *The visual brain in action*. Oxford: Oxford University Press.

Milner, A. D., and Goodale, M. A. (2006) *The visual brain in action*. 2nd edn. Oxford: Oxford University Press.

Mollon, J. D. (1974) After-effects and the brain. *New Scientist*, 61, 479–482.

Mollon, J. D., Pokorny, J., and Knoblauch, K. (2003) *Normal and defective colour vision*. Oxford: Oxford University Press.

Morgan, C. L. (1894) *An introduction to comparative psychology*. London: Walter Scott.

Morgan, M. (2006) *The midbrain*. London: Chelsea House.

Morgan, M. J. (1977) *Molyneux's question. Vision, touch and the philosophy of perception*. Cambridge: Cambridge University Press.

Müller, H. (1851) Zur Histologie der Nethaut. *Zeitschrift für wissenschaftliche Zoologie*, 3, 234–237.

Müller, J. (2003) *Müller's Elements of physiology.* Trans. W. Baly. London: Taylor and Walton. Reprinted, Bristol: Thoemmes.

Munk, H. (1879) Physiologie der Sehsphäre der Grosshirnrinde. *Centralblatt für praktische Augenheilkunde*, 3, 255–266.

Musatti, C. L. (1924) Sui fenomeni stereocinetici. *Archivio Italiano di Psycologia*, 3, 105–120.

Necker, L. A. (1832) Observations on some remarkable phænomena seen in Switzerland: and on an optical phaenomenon which occurs on viewing a crystal or geometrical solid. *Philosophical Magazine and Journal of Science*, 1, 329–337.

Newton, I. (1704) *Opticks: or, a treatise of the reflections, refractions, inflections and colours of light*. London: Smith and Walford.

Ninio, J. (2001) *The science of illusions*. Trans. F. Philip. Ithaca, NY: Cornell University Press.

Ninio, J. (2011) *L'empreinte des sens*. Paris: Odile Jacob.

Ono, H. (1981) On Wells' (1792) law of visual direction. *Perception & Psychophysics*, 32, 201–210.

Ono, H. (1990) Binocular visual directions of an object when seen as single or double. In D. Regan (ed.) *Vision and visual dysfunction, Vol. 9. Binocular vision*. New York: Macmillan. pp. 1–18.

Ono, H., and Wade, N. J. (2005) Depth and motion in historical descriptions of motion parallax. *Perception*, 34, 1263–1273.

Ono, H., and Wade, N. J. (2012) Two historical strands in studying visual direction. *Japanese Psychological Research*, 54, 71–88.

Palmer, S. E. (1999) *Vision science. Photons to phenomenology*. Cambridge, MA: MIT Press.

Palmer, S. E., Schloss, K. B., and Sammartino, J. (2012) Hidden knowledge in aesthetic judgments. In A. P. Shimamura and S. E. Palmer (eds.) *Aesthetic science: Connecting minds, brains, and experience.* New York: Oxford University Press. pp. 189–222.

Panum, P. L. (1858) *Physiologische Untersuchungen über das Sehen mit zwei Augen*. Kiel: Schwerssche Buchhandlung.

Papathomas, T. V. (2007) Art pieces that 'move' in our minds – an explanation of illusory motion based on depth reversal. *Spatial Vision*, 21, 79–95.

Park, D. (1997) *The fire within the eye. A historical essay on the nature and meaning of light*. Princeton NJ: Princeton University Press.

Pelli, D. G., Robson, J. G., and Wilkins, A. J. (1988) The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, 2, 187–199.

Pinto, N., Cox, D. D., and DiCarlo, J. J. (2008) Why is real-world visual object recognition hard? *PLoS Computational Biology*, 4(1): e27. doi:10.1371/journal.pcbi.0040027.

Pirenne, M. H. (1967) *Vision and the eye*. 2nd edn. London: Chapman and Hall.

Pirenne, M. H. (1970) *Optics, painting and photography*. Cambridge: Cambridge University Press.

Plateau, J. (1849) Quatrième note sur de nouvelles applications curieurses de la persistence des impressions de la rétine. *Bulletins de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Bruxelles*, 16, 254–260.

Plater, F. (1583) *De corporis humani structura et usu*. Basel: König.

Polyak, S. (1957) *The vertebrate visual system*, Chicago: University of Chicago Press.

Porta, J. B. (1593) *De refractione. Optices Parte. Libri Novem*. Naples: Carlinum and Pacem.

Porta, J. B. (1669) *Natural Magick*, trans. J. Wright, London.

Porterfield, W. (1759) *A treatise on the eye, the manner and phænomena of vision*. Edinburgh: Hamilton and Balfour.

Priesley, J. (1772) *The history and present state of discoveries relating to vision, light, and colours*. London: Johnson.

Purkinje, J. (1823) *Beobachtungen und Versuche zur Physiologie der Sinne. Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht*. Prague: Calve.

Purkinje, J. (1825) *Beobachtungen und Versuche zur Physiologie der Sinne. Neue Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht*. Berlin: Reimer.

Purves, D., and Lotto, R. B. (2010) *Why we see what we do redux: A wholly empirical theory of vision*. Sunderland, MA: Sinauer Associates.

Ratliff, F. (1965) *Mach bands: Quantitative studies of neural networks in the retina*. San Francisco: Holden-Day.

Rieber, R. W. (ed.) (2012) *Encyclopedia of the history of psychological theories*. New York: Springer.

Rieber, R. W. and Robinson, D. K. (2001) *Wilhelm Wundt in history: The making of scientific psychology*. New York: Springer.

Robinson, A. (2006) *The last man who knew everything: Thomas Young, the anonymous polymath who proved Newton wrong, explained how we see, cured the sick, and deciphered the Rosetta Stone*. New York: Pi Press.

Robinson, D. A. (2008) *Consciousness and mental life*. New York: Columbia University Press.

Robinson, J. O. (1998) *The psychology of visual illusion*. London: Constable.

Rogers, B. (2009) Rogers and Graham's 1979 paper. *Perception*, 38, 907–919.

Rogers, K. (ed.) (2011) *The eye: The physiology of human perception*. New York: Britannica Educational Publishing.

Rolls, E. T. (2011) Face neurons. In A. J. Calder, G. Rhodes, M. H. Johnson, and J. V. Haxby (eds.) *The Oxford handbook of face perception.* Oxford: Oxford University Press. pp. 51–75.

Ronchi, V. (1970) *The nature of light. An historical survey*. Trans. V. Barocas, London: Heinemann.

Roorda, A. (2011) Adaptive optics for studying visual function: A comprehensive review. *Journal of Vision*, 11, 5, 6. doi:10.1167/11.5.6.

Rose, D. (1999) The historical roots of the theories of local signs and labelled lines. *Perception*, 28, 675–685.

Rosenfield, M., Logan, N., and Edwards, K. (eds.) (2009) *Optometry: Science, techniques and clinical management*. 2nd edn. Oxford: Elsevier.

Ross, H. E. and Murray, D. J. (1996) *E. H. Weber on the tactile senses*. Hove, UK: Erlbaum.

Ross, H. E. and Plug, C. (2002) *The mystery of the moon illusion*. Oxford: Oxford University Press.

Ross, W. D. (ed.); (1931) *The works of Aristotle*. Vol. 3. Oxford, UK: Clarendon.

Russell, G. A. (2010) After Galen: Late antiquity and the Islamic world. In F. Boller, S. Finger, and K. Tyler (eds.) *Handbook of clinical neurology. Vol. 95. History of neurology*. 3rd series. Edinburgh: Elsevier. pp. 61–77.

Sabra, A. I. (1989) *The Optics of Ibn Al-Haytham. Books I–III. On direct vision*. London: Warburg Institute.

Sacks, O. (1985) *The man who mistook his wife for a hat*, London: Duckworth.

Sacks, O. (1998) *The island of the colourblind*. London: Vintage.

Sacks, O. (2009) *An anthropologist on Mars.* (4th edn.) London: Picador.

Saito, A. (ed.) (2000) *Bartlett, culture and cognition*, Hove, East Sussex: Psychology Press.

Scheiner, C. (1619) *Oculus, hoc est fundamentum opticum.* Innsbruck: Agricola.

Scheiner, C. (1630) *Rosa Ursina*. Bracciani: Phaeum.

Schenk, T., Franz, V., and Bruno, N. (2011) Vision-for-perception and vision-for-action: Which model is compatible with the available psychophysical and neurophysiological data? *Vision Research*, 51, 812–818.

Schickore, J. (2007) *The microscope and the eye. A history of reflections, 1740–1870*. Chicago: University of Chicago Press.

Schoonover, C. (2010) *Portraits of the mind. Visualizing the brain from antiquity to the 21st century*. New York: Abrams.

Schultze, M. (1866) *Zur Anatomie und Physiologie der Retina*. Bonn: Cohen.

Sekuler, R. and Blake, R. (2005) *Perception*. 5th edn. Maidenhead, UK: McGraw-Hill.

Shapiro, A. and Todorovic, D. (eds.) (2013) *Oxford compendium of visual illusions*. Oxford: Oxford University Press.

Skov, M., and Vartanian, O. (eds.) (2009) *Neuroaesthetics*. Amityville, NY: Baywood.

Slyce, J. (1998) *Patrick Hughes: Perverspective*. London: Momentum.

Smith, A., Heeger, D., Boynton, G., and Norcia, A. (2008) Neuroimaging in vision science. *Journal of Vision*, 8(10), 1.

Smith, A. M. (1996) *Ptolemy's theory of visual perception: An English translation of the* Optics *with introduction and commentary*. Philadelphia: American Philosophical Society.

Snowden, R., Thompson, P., and Troscianko, T. (2012) *Basic vision: An introduction to visual perception*. Oxford: Oxford University Press.

Solomon, J. A. (ed.) (2011) *Fechner's legacy in psychology: 150 years of elementary psychophysics*. Leiden: Brill.

Spillmann, L. (1994) The Hermann grid illusion: A tool for studying human perceptive field organization. *Perception*, 23, 691–708.

Stabell, B., and Stabell, U. (2009) *Duplicity theory of vision. From Newton to the present*. Cambridge: Cambridge University Press.

Stevens, S. S. (1961) To honor Fechner and repeal his law. *Science*, 133, 80–86.

Stevens, S. S. (1975) *Psychophysics*. New York: Wiley.

Stratton, G. M. (1897a). Upright vision and the retinal image. *Psychological Review*, 4, 182–187.

Stratton, G. M. (1897b). Vision without inversion of the retinal image. *Psychological Review*, 4, 341–360.

Stratton, G. M. (1897c). Vision without inversion of the retinal image. *Psychological Review*, 4, 463–481.

Stratton, G. M. (1917) *Theophrastus and the Greek physiological psychology before Aristotle*. New York: Macmillan.

Sutherland, N. S. (1973) Object recognition. In E. C. Carterette and M. P. Friedman (eds.) *Handbook of perception. Vol. 3. Biology of perceptual systems*. New York: Academic Press, pp. 157–185.

Swanston, M. T. (1994) Frames of reference and motion aftereffects. *Perception*, 24, 1257–1264.

Swanston, M. T. and Gogel, W. C. (1986) Perceived size and motion in depth from optical expansion. *Perception & Psychophysics*, 39, 309–326.

Swanston, M. T., and Wade, N. J. (1992) Motion over the retina and the motion aftereffect. *Perception*, 21, 569–582.

Swanston, M. T. and Wade, N. J (1994) A peculiar Optical Phænomenon. *Perception*, 23, 1107–1110.

Swanston, M. T., Wade, N. J., and Day, R. H. (1987) The representation of uniform motion in vision. *Perception*, 16, 143–159.

Swanston, M. T., Wade, N. J. and Ono, H. (1990) The binocular representation of uniform motion. *Perception*, 19, 29–34.

Symons, J, and Calvo, P. (eds.) (2009) *The Routledge companion to philosophy of psychology*. London: Routledge.

Taylor, J. (1738) *Le mechanisme ou le nouveau traité de l'anatomie du globe de l'oeil, avec l'usage de ses différentes parties, & de celles qui lui sont contigues*. Paris: David.

Thirkettle, M., Benton, C. P., and Scott-Samuel, N. E. (2009) Contributions of form, motion and task to biological motion perception. *Journal of Vision*, 9(3): 1–11. doi: 10.1167/9.3.28.

Titchener, E. B. (1910) *A textbook of psychology*. New York: Macmillan.

Turner, G. L'E. (1998) *Scientific instruments 1500–1900. An introduction*. Berkeley, CA: University of California Press.

Turner, R. S. (1994) *In the eye's mind. Vision and the Helmholtz–Hering controversy*. Princeton, NJ: Princeton University Press.

Ungerleider, L. G., and Mishkin, M. (1982) Two cortical visual systems. In D. J. Ingle and M. A. Goodale (eds.) *Analysis of visual behaviour*. Cambridge, MA: MIT Press. pp. 549–586.

Valberg, A. (2005) *Light vision colour*. Chichester, UK: Wiley.

Vernon, M. D. (1970) *Perception through experience*. London: Methuen.

Vesalius, A. (1543) *De humani corporis fabrica Libri septem*. Basel.

Vicario, G. B. (2008) Optical-geometrical illusions: the nomenclature. *Gestalt Theory*, 30, 168–180.

Vicario, G. B.(2011) *Illusioni ottico-geometriche. Una rassegna di problemi*. Venice: Istituto Veneto di Scienze, Lettere ed Arte.

Wade, N. (1982) *The art and science of visual illusions*. London: Routledge & Kegan Paul.

Wade, N. J. (ed.) (1983) *Brewster and Wheatstone on vision*. London: Academic Press.

Wade, N. (1990) *Visual allusions: Pictures of perception*, London: Erlbaum.

Wade, N. J. (1992a). Faces of psychology. *Perception*, 21, Supplement 1. pp. 107ff.

Wade, N. J. (1992b). The representation of orientation in vision. *Australian Journal of Psychology*, 44, 139–145.

Wade, N. J. (1995) *Psychologists in word and image*. Cambridge, MA: MIT Press.

Wade, N. J. (1998a). *A natural history of vision*. Cambridge, MA: MIT Press.

Wade, N. J. (1998b). Early studies of eye dominances. *Laterality*, 3, 97–109.

Wade, N. J. (2003) *Destined for distinguished oblivion: The scientific vision of William Charles Wells (1757–1817)*. New York: Kluwer/Plenum.

Wade, N. J. (2004) Philosophical instruments and toys: Optical devices extending the art of seeing. *Journal of the History of the Neurosciences*, 13, 102–124.

Wade, N. J. (2004a). Joseph Priestley (1733–1804). *Perception*, 33, 509–512.

Wade, N. J. (2004b). Visual neuroscience before the neuron. *Perception*, 33, 869–889.

Wade, N. J. (2005a). *Perception and illusion: Historical perspectives*. New York: Springer.

Wade, N. J. (2005b). Sound and sight: acoustic figures and visual phenomena. *Perception*, 34, 1275–1290.

Wade, N. J. (2007a). Artful visions. *Spatial Vision*, 21, 27–53.

Wade, N. J. (2007b). Image, eye and retina. *Journal of the Optical Society of America A*. 24, 1229–1249.

Wade, N. J. (2008a). Snellen's letters. *Perception*, 37, 167–170.

Wade, N. J. (2008b). Chevalier John Taylor, ophthalmiater. *Perception*, 37, 969–972.

Wade, N. J. (2008c). Visual red (*Sehrot*). *Perception*, 37, 1467–1470.

Wade, N. J. (2008d). Visual purple (*Sehpurpur*). *Perception*, 37, 1617–1620.

Wade, N. J. (2010) Pioneers of eye movement research. *i-Perception*, 1(2), 33–68. http://i-perception.perceptionweb.com/fulltext/i01/i0389.pdf

Wade, N. J. (2011) Eye contricks. *i-Perception*, 2(5), 486–501. http://i-perception.perceptionweb.com/fulltext/i02/i0442aap.pdf

Wade, N. J. (2012a). Perception. In R. W. Rieber (ed.) *Encyclopedia of the History of Psychological Theories*. New York: Springer. pp. 773–788.

Wade, N. J. (2012b). Artistic precursors of Gestalt principles. *Gestalt Theory*, in press.

Wade, N. J., and Brožek, J. (2001) *Purkinje's vision. The dawning of neuroscience*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wade, N. J., and Finger, S. (2001) The eye as an optical instrument. From *camera obscura* to Helmholtz's perspective. *Perception*, 30, 1157–1177.

Wade, N. J., and Gregory, R. L. (2006a). Molyneux's answer I. *Perception*, 35, 1437–1440.

Wade, N. J., and Gregory, R. L. (2006b). Molyneux's answer II. *Perception*, 35, 1579–1582.

Wade, N. J., and Hughes, P. (1999) Fooling the eyes: *trompe l'oeil* and reverse perspective. *Perception*, 28, 1115–1119.

Wade, N. J., and Melcher, D. (2006) Cave art interpretation. I. *Perception*, 35, 577–580.

Wade, N. J., and Nekes, W. (2005) The two faces of Rex Whistler (1905–1944). *Perception*, 34, 639–644.

Wade, N. J., and Ono, H. (2012) Early studies of binocular and stereoscopic vision. *Japanese Psychological Research*, 54, 54–70.

Wade, N. J., and Piccolino, M. (2009) Seeing stars. *Perception*, 38, 1581–1584.

Wade, N. J., and Swanston, M. T. (1987) The representation of nonuniform motion: induced movement. *Perception*, 16, 555–571.

Wade, N. J., and Swanston, M. T. (1996) A general model for the perception of space and motion. *Perception*, 25, 187–194.

Wade, N. J., and Swanston, M. T. (2001) Helmholtz on golf. *Perception*, 30, 1407–1410.

Wade, N. J., and Tatler, B. W. (2005) *The moving tablet of the eye: The origins of modern eye movement research*. Oxford: Oxford University Press.

Wade, N. J., Campbell, R. N., Ross, H. E., and Lingelbach, B. (2010) Necker in Scotch Perspective. *Perception*, 39, 1–4.

Wade, N. J., Kovács, G., and Vidnyánszky, Z. (2003) Inverted faces. *Perception*, 32, 1–6.

Wade, N. J., Spillmann, L. and Swanston, M. T. (1996) Visual motion aftereffects: critical adaptation and test conditions. *Perception*, 36, 2167–2175.

Wade, N. J., Swanston, M. T., and de Weert, C. M. M. (1993) On interocular transfer of the motion aftereffect, *Perception*, 22, 1365–1380.

Wallach, H. and O'Connell, D. N. (1953) The kinetic depth effect. *Journal of Experimental Psychology*, 45, 205–217.

Walsh, V., and Kulikowski, J. J. (1998) *Perceptual constancy: Why things look as they do.* Cambridge: Cambridge University Press.

Walsh, V. and Pascual-Leone, A. (2003) *Transcranial magnetic stimulation: a neurochronometrics of mind.* Cambridge, MA: MIT Press.

Wandell, B. (1995) *Foundations of vision*. Sunderland, MA: Sinauer Associates.

Wandell, B. A., Brewer, A. A., and Dougherty, R. F. (2005) Visual field map clusters in human cortex. *Philosophical Transactions of the Royal Society. Series B.* 360, 693–707.

Watson, J. B. (1913) Psychology as the behaviorist views it. *Psychological Review*, 20, 158–177.

Weiskrantz, L. (1986) *Blindsight. A case study and implications*. Oxford: Oxford University Press.

Weiskrantz, L. (2009) *Blindsight. A case study spanning 35 years and new developments.* 2nd edn. Oxford: Oxford University Press.

Wells, W. C. (1792) *An essay upon single vision with two eyes: together with experiments and observations on several other subjects in optics*. London: Cadell.

Werner, J. S. (1998) Aging through the eyes of Monet. In W. G. K. Backhaus, R. Kliegl, and J. S. Werner (eds.) *Colour vision. Perspectives from different disciplines*. Berlin: de Gruyter. pp. 1–41.

Werner, J. S., Pinna, B., and Spillmann, L. (2007) The brain and the world of illusory colour. *Scientific American*, 293(3), 90–95.

Wertheimer, M. (1922) Untersuchungen zur Lehre von der Gestalt. I. Prinzipielle Bemerkungen. *Psychologische Forschung*, 1, 47–58.

Wertheimer, M. (1923) Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung*, 4, 301–350.

Westheimer, G. (2008) Was Helmholtz a Bayesian? *Perception*, 37, 642–650.

Westheimer, G. (2011) Reversed tilt effect for dichoptic stimulation in vertical meridian. *Vision Research*, 51, 101–104.

Westwood, D. A., and Goodale, M. A. (2011) Converging evidence for diverging pathways: Neuropsychology and psychophysics tell the same story. *Vision Research*, 51, 804–811.

Wheatstone, C. (1838) Contributions to the physiology of vision – Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society*, 128, 371–394.

Whitaker, H. A., and Turgeon, Y. (2008) Charles Bonnet's neurophilosophy. In H. Whitaker, C. U. M. Smith, and S. Finger (eds.) *Brain, mind, and medicine: Essays in 18th century neuroscience*. New York: Springer. pp. 191–200.

Willatts, J. (1997) *Art and representation: New principles in the analysis of pictures.* Princeton, NJ: University of Princeton Press.

Wundt, W. (1874) *Grundzüge der physiologischen Psychologie*. Leipzig: Engelmann.

Yannuzzi, L. A. (2010) *The retinal atlas*. Oxford: Elsevier Saunders.

Young, A. (2011) Disorders of face perception. In A. J. Calder, G. Rhodes, M. H. Johnson, and J. V. Haxby (eds). *The Oxford handbook of face perception.* Oxford: Oxford University Press. pp. 77–91.

Young, L. R., Henn, V., and Scherberger, H. (2001) *Fundamentals of the theory of movement perception by Dr Ernst Mach*. Dordrecht: Kluwer/Plenum.

Young, T. (1800) Outline of experiments and enquiries respecting sound and light. *Philosophical Transactions of the Royal Society*, 90, 106–150.

Young, T. (1802) On the theory of lights and colours. *Philosophical Transactions of the Royal Society*, 92, 12–48.

Young, T. (1807) *A course of lectures on natural philosophy and the mechanical arts*. London: Johnson.

Young, T. (2002) *Thomas Young's Lectures on natural philosophy and the mechanical arts*. Bristol: Thoemmes Press.

Zanforlin, M. (2006), Illusory space and paradoxical transparency in stereokinetic objects. In L. Albertazzi (ed.) *Visual thought: the depictive space of perception*. Amsterdam: Benjamins. pp. 99–104.